MEASURING THE RESULTS OF TEACHING

MONROE

RIVERSIDE TEXTBOOKS
IN EDUCATION
EDITED BY ELLWOOD P. CUBBERLEY
PROFESSOR OF EDUCATION
LELAND STANFORD JUNIOR UNIVERSITY

DIVISION OF SECONDARY EDUCATION
UNDER THE EDITORIAL DIRECTION
OF ALEXANDER INGLIS
ASSISTANT PROFESSOR OF EDUCATION
HARVARD UNIVERSITY

Digitized by the Internet Archive
in 2007 with funding from
Microsoft Corporation

http://www.archive.org/details/measuringresults00monruoft

Educational
Teaching
M.

MEASURING THE RESULTS OF TEACHING

BY

WALTER SCOTT MONROE, Ph.D.
DIRECTOR OF BUREAU OF COÖPERATIVE RESEARCH
SCHOOL OF EDUCATION, INDIANA UNIVERSITY

150977
5-6-19

HOUGHTON MIFFLIN COMPANY
BOSTON NEW YORK CHICAGO
The Riverside Press Cambridge

COPYRIGHT, 1918, BY WALTER SCOTT MONROE

ALL RIGHTS RESERVED

The Riverside Press
CAMBRIDGE . MASSACHUSETTS
U. S. A.

EDITOR'S INTRODUCTION

Up to very recently the quality of instruction given by a teacher has been almost entirely a matter of personal opinion. A teacher was good, average, or poor, largely as he or she impressed a principal or supervisor; and the basis for estimating these qualities lay largely in the theory as to the nature of the educational process possessed by the supervisory officer. To a superintendent of the drill and memorization and martinet type, a creative and stimulative and original teacher probably would be classed as poor; whereas such a teacher would be highly prized by a superintendent in close sympathy with the creative and expressive tendencies in modern education. Against such personal opinion the teacher has had almost no means of defense. On the other hand, it has been almost equally difficult for a superintendent to demonstrate to questioning laymen wherein the work of a teacher lacked effectiveness.

Within recent years a number of personal-estimate scales have been devised by students and superintending officers for charting, in visual form, the important characteristics of teachers. Such charts have been useful in revealing points of strength and weakness, both to teachers and to supervisory officers.

Practically within the past five years an entirely new series of instruments for estimating teaching efficiency has been made available in the form of the new Standardized Tests, with their accompanying Standard Scores and Score Charts. It is with this new set of measuring tools that this volume deals. Much of the early work in evolving and standardizing these new measuring scales, and accumulating results from which to work out the Standard Scores, naturally had to be quite technical and was hard for the teacher to understand. Enough such work has now been done to enable the author of this volume to organize and present, in simple and readable form, the essential information needed by grade teachers to enable them to use these Standardized Tests to measure and determine for themselves the effectiveness of their own instruction, what are the points of strength and weakness in the work they are doing, and where they should add emphasis and where enough emphasis has been placed. The years of important work done by the author in directing the teachers of the State of Kansas in estimating and evaluating the work of the Kansas schools should in itself insure a helpful volume.

The value of such a book as this one to the teacher in service cannot but be large. It is seldom that books of such definiteness are written for the use of teachers. A study and mastery of the method of this volume mean the acquirement of a new tool for estimating personal efficiency and self-improvement. The use of the Tests means a new ability to diagnose and prescribe. To the work of the teacher in the classroom they give a definiteness heretofore unknown. To use a military term, they set the "limited objectives" for each subject of the course of study, which the teacher is expected to reach, but beyond which she is not expected to go. They prevent a waste of teaching energy by preventing over-emphasis, and set standards in instruction which are indisputable because they are based on the school practice of the best schools of the United States. By their use teachers may determine their own efficiency, compare the progress of their pupils or class, with pupils or classes elsewhere in terms that are definite and measures that are comparable; and, if unjustly criticized, they can defend the work they are doing. The new Standardized Tests give a definiteness and scientific accuracy to the work of schoolroom instruction heretofore unknown, and teachers in all kinds of school systems will be benefited by a careful study of this important volume.

Ellwood P. Cubberley

PREFACE

This book is written for the teacher in the elementary school. As such it is not intended to be a fundamental treatise upon educational measurements, but rather a text which will help teachers to use standardized tests to the greatest advantage. Only certain ones of the available standardized tests are described. This was thought to be a more helpful plan than to include all of the available tests, because the elementary teacher seldom has at hand the information necessary for an intelligent selection of standardized tests. The value of a number of the tests described in this text has been demonstrated by wide usage. Others are just being made available for use. In these cases it has been necessary for the author to exercise his judgment based upon four years' experience in supplying tests to teachers and superintendents through the Bureau of Educational Measurements and Standards of the Kansas State Normal School, Emporia, Kansas. Some worthy tests have been omitted partly because of the limitations of space and partly because of other considerations. Other worthy tests will doubtless be devised in the future, some of them replacing certain of the tests chosen for description in this text.

The feasibility of the test being used by teachers who have not had special training in the field of educational measurements has been kept constantly in mind. Detailed directions for the use of a test are generally not reproduced, for they are furnished with the test when purchased for class use. Only those general features which are necessary for understanding the tests and the method of handling the results have been given. It is believed that any teacher who studies carefully the descriptions given in this text will have no difficulty in using any of the tests described.

It is the contention of the author that the use of a standardized test is justified only when the teacher can use the resulting measures as a basis for improving instruction. Consequently much space is given to the interpretation of scores or measures and the corrective instruction which should be given to correct unsatisfactory scores. Unfortunately little is known about corrective measures for certain school subjects. This, however, is a condition which time will remedy.

The author is aware that in a sense he has done little more than bring together the results of a number of workers in this field and he realizes his indebtedness to them. He is particularly indebted to Dean F. J. Kelly and Captain J. C. DeVoss who kindly permitted him to use portions of their chapters in *Educational Tests and Measurements.*

WALTER S. MONROE

BLOOMINGTON, INDIANA

CONTENTS

I. The Inaccuracy of Present School Marks . . . 1

II. The Measurement of Ability in Reading 22

III. The Meaning of Scores and Correcting Defects in Reading 43

IV. The Measurement of Ability in the Operations of Arithmetic 97

V. Diagnosis and Corrective Instruction in Arithmetic 118

VI. The Measurement of Ability to solve Problems and Corrective Instruction 154

VII. The Measurement of Ability in Spelling and Corrective Instruction 175

VIII. The Measurement of Ability in Handwriting . 203

IX. The Measurement of Ability in Language and Grammar 235

X. The Measurement of Ability in Geography and History 255

XI. Educational Measurements and the Teacher . 267

XII. Summary 281

Appendix 285

Index 295

LIST OF FIGURES

1. Distribution of measures of silent reading ability as measured by the Kansas Silent Reading Tests 5
2. Distribution of marks in University of Chicago High School in English and History; and distribution of marks of two teachers in the same department 7
3. Distribution of marks assigned to one geometry paper by 116 teachers 9
4. Form used in recording the scores obtained by using Monroe's Standardized Silent Reading Tests 28
5. Forms used in recording the scores obtained by using Courtis's Silent Reading Test No. 2 34
6. Record Sheet for recording scores obtained by using Thorndike's Visual Vocabulary Scale 38
7. A scheme for the graphical representation of the scores of Monroe's Standardized Silent Reading Tests; and the scores of four seventh-grade pupils 45
8. Median scores of a school in silent reading as determined by the Courtis Silent Reading Test No. 2 48
9. A scheme for the graphical representation of scores on Gray's Oral Reading Test 50
10. Scores of a fifth-grade class on Monroe's Standardized Silent Reading Tests. Type I 52
11. Scores of a fourth-grade class on Monroe's Standardized Silent Reading Tests. Type II 64
12. Per cent of 1831 Cleveland pupils found in each on nine speed and quality groups in silent reading 68
13. Scores of a fifth-grade class on Monroe's Standardized Silent Reading Tests. Type III 70
14. Silent reading rates of fourth-grade pupils, showing effect of corrective treatment 81
15. Average number of pages of silent reading per pupil during school hours. Supplementary material 82

16. Scores of a 2 A class on the Courtis Silent Reading Test No. 2. Type IV 83
17. Scores of a sixth-grade class on the Courtis Silent Reading Test No. 2. Type V 85
18. Improvement in oral-reading rate of twenty-fourth when individual and group instruction was used 89
19. Form of tabulation sheet for recording scores obtained by using the Courtis Standard Research Tests, Series B, and the scores of a seventh-grade class in addition . . . 101
20. Use of Courtis's Graph Sheet No. 3 110
21. Type I. The record sheet of a fifth-grade class in addition 120
22. Type IV. The record sheet of an eighth-grade class in multiplication 125
23. Median scores of two classes in the same city . . . 129
24. Two records of one girl 130
25. Median scores of three sixth-grade classes 132
26. Individual scores of three sixth-grade pupils in the same class. (Class A in Fig. 25) 134
27. Two records of one pupil on the Cleveland Survey Tests 152
28. Distribution of 91 pupils according to the number of words spelled correctly 179
29. Record sheet for recording pupils' scores on a spelling test of fifty words 188
30. Form of class record sheet for recording scores in handwriting 213
31. Individual record card, Freeman Scale 216
32. Standard Score Card for measuring handwriting . . . 217
33. Graphical representation of Ayres's Standards for the "Gettysburg Edition" of his Handwriting Scale . 220
34. Distribution of scores in handwriting of a third-grade class 224
35. Distribution of scores in handwriting of a fourth-grade class 230
36. Distribution of scores in handwriting of a fifth-grade class 231
37. Standard distribution of scores in handwriting . . . 232

38. Class record sheet for use with Willing's Composition Scale 242
39. Record sheet for diagnosis. Charters's Diagnostic Test in Language and Grammar 247
40. Illustrating Courtis's Standard Test in Geography for States and important cities of the United States . . . 257
41. A section of the Hahn-Lackey Geography Scale . 260, 261
42. Medians in speed and accuracy in addition test for pupils of grades 4 to 8 inclusive, showing Courtis's medians, medians for Cuyahoga County, and for three districts in the county 270
43. Median scores of a sixth-grade class in September, 1917, and in April, 1918, as measured by the Courtis Standard Research Tests in Arithmetic, Series B 273
44. Improvement in spelling "efficiency" in certain schools of Cleveland, Ohio 275
45. Effect of continuous use of the Courtis Standard Research Tests, Series B, in Boston, Eighth Grade, 1915 . 276

LIST OF TABLES

1. The rank of problems by teachers' judgments . . . 13
2. Problems ranked according to real difficulty and teachers' estimates 14
3. A poor arrangement of scores 27
4. The same scores rearranged in a better order . . . 27
5. Scores of seven seventh-grade pupils on Monroe's Standardized Silent Reading Tests 43
6. Standard May scores for Monroe's Standardized Silent Reading Tests 44
7. Standard scores for Courtis's Silent Reading Test No. 2 46
8. The scores of one school 49
9. Median scores in visual vocabulary (Thorndike Scale A) 49
10. Standard scores for Gray's Oral Reading Test . . . 51
11. Rate of silent reading in informal testing 67
12. A typical distribution of scores in number of examples attempted 103
13. Three special cases which arise in using Courtis's Standard Research Tests, Series B 106
14. Standard median scores, Courtis's Standard Research Tests, Series B 108
15. The distribution of the pupils of a city according to the number of examples attempted, Courtis's Standard Research Tests, Series B 126
16. The range of number of examples attempted . . . 127
17. Frequency of types of errors in subtraction, multiplication, and division based upon a study of 812 test papers, Courtis's Standard Research Tests, Series B . . . 143
18. Percentage of failures on vocabulary test in arithmetic 166
19. Percentage of pupils who failed to draw correctly the figures named 167
20. Median scores for a timed-sentence spelling test of fifty words 190

21. The misspelling of eighty seventh-grade pupils on a column spelling test 195
22. Handwriting standards. Rate in letters per minute. Quality in terms of Ayres's Scale 219
23. Median scores for Willing's Composition Scale . . 243
24. Distributions of differences between two teachers' marks on sets of fifth-grade arithmetic papers — first, without any effort to unify the methods used, and second, by a common standard 279

MEASURING THE RESULTS OF TEACHING

CHAPTER I

THE INACCURACY OF PRESENT SCHOOL MARKS

The measurement of results not new in education. Educational measurements are not new in school work, although this name has not been applied to them until very recently. Since schools have existed, teachers and other school officials have attempted to measure the abilities of pupils by estimating daily recitations and by examinations. The measures of the abilities of pupils obtained in these ways are thought to possess a high degree of precision and are considered very important.

The promotion of pupils depends upon the "grades" they receive. The ability of a pupil in each of the subjects is measured by the teacher's estimate and by examination, and if the resulting measures show the pupil to be a few points, or in some instances a fraction of a point below the "passing mark," the pupil is classified as a failure. If the resulting measures equal or are above the "passing mark," the pupil is promoted.

The "grades" or school marks are entered upon the monthly or quarterly report cards. Parents, as well as teachers and pupils, take these school marks very seriously. If Johnnie's "grades" for a given month are below those of the preceding months, or, worse still, if they are below those of neighbor Smith's Mary, an explanation is demanded. A permanent record is kept of at least the yearly "grades," and the awarding of school honors is based upon it.

Until recently, practically all admission to college was determined by examination. Except in the universities and colleges of the Central and Western American States, the custom still maintains generally throughout the world. This practice is based on the assumption that the examining committee can determine thereby the effectiveness of the candidate's college preparatory work. The civil service, from its inception in China centuries ago until the present day, has employed the examination as a means for measuring the ability of persons who desire positions operated under this system.

The use of scientific tests and standards is new. Although the measurement of the results of instruction is not new, it should be recognized that the use of tests which have been scientifically constructed and the interpretation of the resulting measures by comparison with standards is one of the most recent educational developments. Thorndike, "the father of this movement," has discovered what is probably the earliest record of this new type of educational measurement. The date of this publication, which was by an English schoolmaster, is 1864. In our own country, Rice's report on spelling in 1897 marked the beginning of the movement, but except for this and a few other pioneer efforts, the development has been confined to the last ten years. Within this period those who ridiculed the work of Rice have been converted to educational measurements, and standardized tests are now generally recognized as one of the most helpful instruments at the command of the teacher and the supervisor.

Recent investigations have shown school marks to be inaccurate. One of the most important factors contributing to our present use of standardized tests has been a number of investigations made to ascertain the accuracy or reliability of measures obtained by means of teachers' estimates and by means of examinations. In the world of physical things we measure distance by means of the yardstick, mass by means of scales, the volume of liquids by means of gallon measures. Measurements of these magnitudes, when made carefully with accurate instruments, possess a high degree of reliability. By a high degree of reliability we mean, for example, that if two persons measure the length of the same room by means of the same yardstick or any other yardstick, the two measurements will be approximately equal. If they differ by more than one or two inches, we doubt the accuracy of both, and we demand that the room be measured again. Similarly, in the case of school-children, if we find that, when the same children are measured in the same subjects by two different teachers, the two sets of measures do not agree rather closely, we have reason to doubt the accuracy of both sets of measures. On the other hand, if the two sets of measures ("grades") agree closely, we have reason to believe them accurate or reliable.

In this chapter we present evidence from three types of investigations which show that marks given by teachers under ordinary conditions are not accurate measures of the abilities of their pupils: (1) Kelly's investigation based upon the final "grades" given to pupils in two successive years by different teachers; (2) Johnson's investigation based upon the distribution of "grades"; (3) the marking of examination papers.

(*1*) *Kelly's investigation.* In 1913, Kelly[1] made an investigation of the marks given to the sixth-grade pupils in four ward schools in Hackensack, New Jersey, and the marks given to the same pupils when they went to a common departmental school for seventh-grade work. This will be recognized as a case where the abilities of the same pupils were measured by two different sets of teachers, the sixth-grade teachers in the ward schools and the seventh-grade teachers in the departmental school. Since in the departmental school all of the pupils were taught arithmetic by one teacher, there was an opportunity to compare the "grades" given in arithmetic by the sixth-grade teachers in the different ward schools. If these teachers were accurate in their "grading," we would expect to find that all of the pupils who received a mark of "G" (good) in arithmetic in the sixth grade would receive approximately the same mark in the seventh grade. If, however, the sixth-grade teachers were inaccurate in their marking, — that is, some of them marked too high or too low, — we would expect to find that pupils having the mark of "G" in the sixth grade, but coming from different schools, would, on the average, receive different marks in the seventh grade. This condition was found to exist.

---

[1] Kelly, F. J., *Teachers' Marks.* (Teachers College Contributions to Education, no. 66, p. 7.)

Kelly states his conclusions as follows:

> This means that for work which the teacher in school "C" (one of the ward schools) would give a mark of "G" (good) in language, penmanship, or history, the teacher in school "D" (another ward school) would give less than a mark "F" (fair).

(*2*) *Johnson's investigation.* Another type of investigation has been made by Johnson,[1] Principal of the University High School of the University of Chicago. It is based upon the fact that when accurate measurements are made of any ability of a large group of pupils, the resulting measures are distributed; that is, arranged along the scale of measurement, in a certain definite way. For example, in Fig. 1 there are represented graphically four distributions of the measures of silent reading ability secured by giving the Kansas Silent Reading Tests. The number of measures represented in each grade is over 5000. The base line of the curve in each

---

[1] Johnson, F. W., "A Study of High School Grades"; in *School Review*, vol. 19, pp. 13–24. See also Kelly, F. J., *Teachers' Marks*, p. 11, and following, for reports of similar investigations.

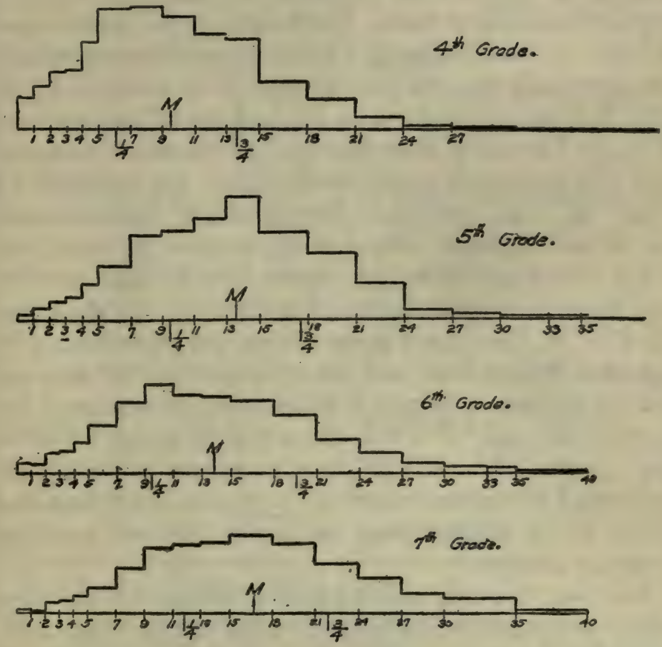


FIG. 1. SHOWING THE DISTRIBUTION OF MEASURES OF SILENT READING ABILITY AS MEASURED BY THE KANSAS SILENT READING TESTS.

case represents the scale of the test, 0, 1, 2, 3, 4, 5, and so on. At any point of this base line the height of the broken line curve above the base line represents the number of pupils having the measure represented. The general shape of these four broken line curves is the same. A few pupils received very low measures and a few very high ones. The great majority of the measures are grouped near the middle where the curve is highest. A curve which, beginning with the low measures, rises gradually and then falls gradually, as do those shown in Fig. 1, is called a "normal curve" and represents the shape of the distribution when accurate measurements have been made. If the shape of the curve representing the distribution of a particular set of measures differs materially from the general shape of the curves in Fig. 1, there is reason for questioning the accuracy of the measures.

In the University High School, "F" denotes failure, and the four successive ranks above failure are indicated by "D," "C," "B," and "A." For the several departments of the school, Johnson tabulated the number of times each mark was given during the years 1907–08 and 1908–09. The conditions which he found to exist may be illustrated by Fig. 2. The upper figure shows the distributions of marks in English (left) and history (right). It will be noted that in the case of English a much larger proportion of low marks ("F" and "D") were given than in history. For the high marks ("A" and "B") just the reverse is true. Both curves fail to conform closely to the normal curve described above which suggests that the marks may not represent accurate measures.

However, the most striking part of the figure is the lower which represents the distributions of the marks of two teachers in the same department. The distribution for teacher A conforms reasonably close to the normal curve, but that for teacher B departs from it in a very conspicuous fashion. It is obvious that teacher B is accustomed to give "high grades." In so doing he has furnished evidence that his marks are probably inaccurate.

(3) *Marking examination papers.* The written examination is the most common means of measuring the abilities of pupils, although many teachers and school patrons oppose

its use. They contend that pupils working under pressure frequently become nervous and confused and consequently cannot do themselves justice, while other pupils, who have no real grasp of the subject, are able by cramming to write excellent papers. It is also contended that the questions are frequently not well selected and do not pertain to the essentials of the subject.

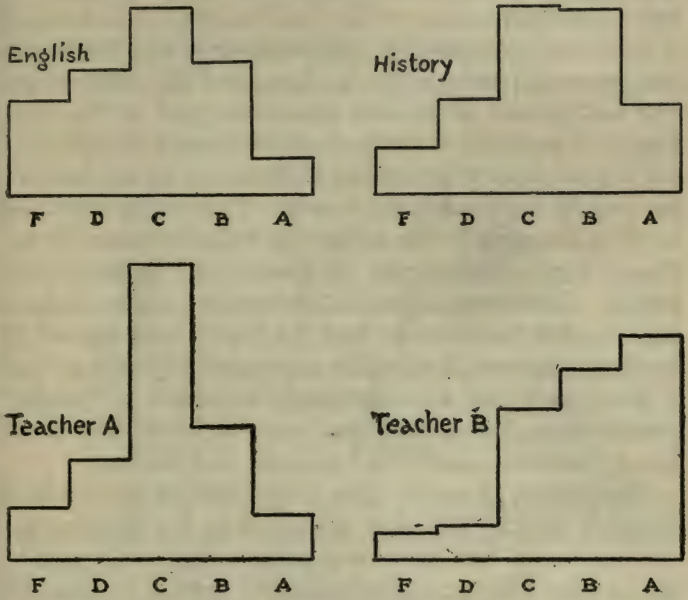


Fig. 2. (*Upper.*) Showing Distribution Marks in University of Chicago High School in English and History. (*Lower.*) Showing Distribution of Marks of Two Teachers in the Same Department. (*After Johnson.*)

There is probably some truth in the above assertions, but within the past few years there have been a number of investigations to ascertain if teachers mark examination papers accurately, assuming that what appears on the papers is a true record of the abilities of the pupils. Starch and Elliott[1] investigated the accuracy with which teachers marked papers in English, geometry, and history. Their method and the facts revealed may be illustrated by the case of geometry.

A facsimile reproduction was made of an actual examination paper in plane geometry. A copy of this reproduction was sent to each of the high schools included in the North Central Association of Colleges and Secondary Schools, with the request that it be marked on the scale of one hundred per cent by the teacher of geometry. The teacher was asked to mark the paper by the method he was accustomed to use. Papers were returned from 116 schools, and the results tabulated. When we consider that the subject-matter of geometry is quite definite, and that the papers were marked by teachers who were thoroughly acquainted with the subject, it would seem that we might expect the marks or "grades" placed upon this examination paper to be in close agreement. However, exactly the opposite was the case.

Distribution of marks. The distribution of the marks is shown in Fig. 3. The scale is marked on the base line and the number of dots above any point indicates the number of teachers who gave the indicated "grade." Thus the "grade" of 75 was given by thirteen teachers, the "grade" of 76 by three teachers, and so on. Of the 116 marks, two were above 90, while one was below 30. Twenty were 80 or above, while twenty other marks were below 60. Forty-seven teachers assigned a mark passing or above, while sixty-nine teachers thought the paper not worthy of a passing mark.

---

[1] Starch and Elliott, "Reliability of Grading High-School Work in English"; in *School Review*, vol. 20, pp. 442–57; "Reliability of Grading Work in Mathematics"; in *School Review*, vol. 21, pp. 254–59; "Reliability of Grading Work in History"; in *School Review*, vol. 21, pp. 676–81.

Not only were similar results obtained by Starch and Elliott in English and in history, but other investigators[1] have verified them many times. In the face of such facts only one conclusion is possible; namely, that under ordinary conditions the marks assigned to examination papers by teachers are very unreliable. Such marks can represent only very crude and very inaccurate measures of the abilities of pupils. It is not too much to say that the mark which a pupil receives on an examination paper depends upon the teacher who "grades" the paper, as well as upon what the pupil places upon the paper.

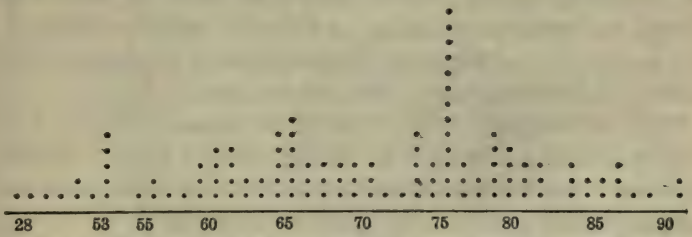


FIG. 3. DISTRIBUTION OF MARKS ASSIGNED TO ONE GEOMETRY PAPER BY 116 TEACHERS.

Passing grade 75. Range 28 to 92. Marks assigned by schools whose passing grade was 70 were weighted by 3 points. Median 70. Probable 7.5.

It has also been shown that the same teacher is not consistent in his own marking. If a set of papers are marked a second time, the two sets of marks will vary widely.[2]

Summary. We have now presented an illustration of each of three types of evidence that teachers' marks, both final "grades" and examination "grades," are inaccurate. In each case the illustration is typical of a number of similar ones which might be mentioned. We have, therefore, a large amount of evidence that teachers' marks are not accurate. We shall next consider two causes for the errors in marking examination papers.

---

[1] See Kelly, F. J., *Teachers' Marks*, p. 51, and following, for accounts of other investigations.

[2] See Starch, Daniel, *Educational Measurements*, p. 9.

Conditions which contribute to the inaccuracy in marking examination papers. (*1*) *Error due to unequal value of questions.* A critical study of examinations and of the manner of giving them reveals certain conditions which contribute to the inaccuracy in teachers' marks. In the first place, the questions are generally considered equal in value, but if we judge the value of questions on the basis of their difficulty as shown by the responses of the pupils, it is seldom that the same credit should be given for answering correctly two different questions. As evidence of this consider the following questions taken from an examination in United States history.[1] The number following the question is the per cent of pupils who answered the question correctly.

To what religious body did most of the settlers of Pennsylvania belong?	62.3
What critical problem arose during Buchanan's administration?	7.0
What is the main purpose of the Monroe Doctrine?	25.5

These differences in the per cent of correct answers are merely typical of what is very likely to be the case in any examination prepared by the teacher. The questions will not be equally difficult, and it is the general practice to base the credit given for a correct answer upon the difficulty of the question: that is, less credit is given for answering correctly an "easy" question than for a "hard" one.

It is easy to understand how a serious element of error is introduced when each question is considered to have a value of ten points and the questions are not equal in difficulty.

---

[1] Buckingham, B. R., "Survey of the Gary and Prevocational Schools," *Seventeenth Annual Report of the City Superintendent of Schools* (New York City), 1914–15.

The situation is much the same as we should have in measuring distances if yardsticks of different lengths were used, but were considered to be equal. Under such circumstances a yard would have no definite length, and to say that a certain distance was 21.42 yards would convey no definite information about it. For this reason the Federal Government has standardized all weights and measures by establishing definite units, and before we can obtain definite measures of the abilities of children, it will be necessary to devise tests consisting of standard units: that is, the questions or exercises composing the test must be evaluated.

A teacher's estimate of the difficulty of questions is unreliable. Can a teacher judge of the difficulty of a problem or even arrange a list of problems in order of difficulty? One investigator[1] studied this question by submitting the following list of twenty-three problems to twenty teachers who were asked to estimate the per cent of pupils who would solve each problem correctly if given ten minutes for each. From this information it was possible to determine which problem each teacher considered easiest, which second in difficulty, and so on. The results of these teachers' judgments are given in Table I.

1. How much change should I expect from $5, after paying for 5 pounds of coffee at 38 cents a pound?
2. If $1991 a day is paid to 724 men who each earn the same wages, how much does each man receive?
3. A boy had 210 marbles. He lost 1/3 of them. How many were left?
4. A grocer had a tank holding 44 3/16 gallons of oil. One day he drew out 15 3/4 gallons and the next day 9 1/8 gallons. How many gallons were left in the tank?
5. There are 550 pupils on the roll. If 5/8 of them are here today, how many are absent?

---

[1] Comin, Robert, "Teachers' Estimates of the Ability of Pupils"; in *School and Society*, vol. 3, p. 67, January 8, 1916.

6. If 3/4 of a pound of cheese is sold for 45 cents, how much can be bought for $1?

7. A storekeeper sold 12 yards of cloth, which was 4/15 of the whole piece. How many yards in the whole piece?

8. A baseball team played 160 games during the season and won 100 of them. What part of the whole number of games did the team win?

9. A store takes in the following sums: $1250.50, $300, $175, $16.25, $120.50, $32.75, $68.50. It pays out: $600, $360, $166.67, $33.33, $240. How much remains after payments are made?

10. A man bought a house for $7250. After spending $321.50 for repairs, he sold it for $9125. How much did he gain?

11. A reader has 29 lines on a page and in all 10,034 lines. How many pages in the book?

12. A boy lost one fourth of his kite string in a tree, one third in some wire, and one fifth in a hedge. What part of his string was left?

13. How much will 8 3/4 dozen pencils cost at the rate of $1/4 for half a dozen?

14. If it takes a train three quarters of an hour to reach a certain station, what fraction of an hour will it take the train to go 3/5 of the distance?

15. A man has a salary of $125 a month. He saves 20 per cent of his salary. How much will he save in a year?

16. A workman pays $22 a month for board, which is 20 per cent of his wages. What are his wages?

17. Mr. Marshall receives a salary of $2500 a year. His rent costs him 1/3 of this and his other expenses are $1500. He saves the rest. What per cent of his salary does he save?

18. John had $1.20 Monday. He earned 30 cents each day on Tuesday, Wednesday, Thursday, and Friday. Saturday morning he spent one third of what he had earned in the four days. Saturday afternoon his father gave John half as much as John then had. How much did his father give John?

19. A boy had $3. He paid it all for four articles, which we will call A, B, C, and D. B cost as much as D. A cost as much as B, C, and D together. The boy sold A and B for 1 1/2 times what he paid for them. He sold C and D for 1 1/4 times what he paid for them. How much did he get for the four articles?

20. A party of children went from a school to a woods to gather nuts. The number found was but 205, so they bought 1955 nuts

more from a farmer. The nuts were shared equally by the children and each received 45. How many children were there in the party?

21. One summer a farmer hired 43 boys to work in an apple orchard. There were 35 trees loaded with fruit and in 57 minutes each boy had picked 49 apples. If in the beginning the total number of apples on the trees was 19,677, how many were there still to be picked?

22. A girl found that by careful counting there were 87 letters more on a page of her history than on a page of her reader. She read 31 pages in each book in the first 29 days of school. How many more letters each day did she read in one book than in the other?

23. The children of a school made small boxes to be filled with candy and given as presents at a school party. Six hundred boxes were needed. In 4 days grades 3 to 7 made 20, 25, 83, 150, and 150 boxes. The eight grade agreed to make the rest. How many did the eighth grade make?

TABLE I. SHOWING THE RANK OF PROBLEMS BY TEACHERS' JUDGMENTS

Rank	*Problems*																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	4	2	12	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
2	5	..	4	..	1	..	..	3	2	1	4	..	..	..	..	..	..	..	..	..	..	..	..
3	2	4	..	..	6	..	2	2	1	1	2	..	..	..	..	..	..	..	..	..	..	..	..
4	2	3	1	2	1	..	2	..	2	..	4	1	..	..	..	2	..	..	..	..	..	..	..
5	2	4	1	..	..	1	..	4	1	..	2	..	..	..	1	3	..	..	..	..	..	..	1
6	1	3	1	1	3	..	1	2	2	..	3	..	..	..	..	..	..	..	..	1	..	..	2
7	2	2	1	1	1	..	..	1	3	4	1	1	..	..	..	1	..	..	..	2	..	..	..
8	2	1	..	3	2	..	1	1	2	3	1	1	..	..	1	..	1	..	..	..	..	..	1
9	..	..	..	2	3	..	1	1	2	2	1	..	1	..	3	3	..	..	..	1	..	..	..
10	..	1	..	3	1	..	2	2	1	1	..	..	1	2	3	1	1	1	..	..	..	..	..
11	..	..	..	2	1	..	3	..	..	..	1	5	2	1	2	2	..	..	..	..	..	..	1
12	..	..	..	1	1	..	3	..	1	2	1	2	1	..	3	2	1	..	..	1	..	..	1
13	..	..	..	1	..	1	2	2	1	..	..	3	4	1	2	1	1	..	..	..	..	..	1
14	..	..	..	1	..	3	1	..	1	4	..	..	2	1	1	1	1	..	..	3	..	..	1
15	..	..	..	1	..	1	1	..	..	1	..	1	3	2	1	1	3	1	..	2	..	1	1
16	..	..	..	..	..	4	..	1	..	..	..	3	..	5	1	1	1	2	..	1	..	1	..
17	..	..	..	1	..	4	..	..	..	..	..	3	3	2	1	1	3	..	..	..	..	..	2
18	..	..	..	1	..	1	..	..	1	..	..	..	..	1	1	1	..	3	..	6	1	..	4
19	..	..	..	..	..	1	..	..	..	1	..	..	2	1	..	..	..	2	1	1	3	5	3
20	..	..	..	..	..	3	1	..	..	..	..	..	..	..	..	..	4	3	..	1	3	5	..
21	..	..	..	..	..	..	..	..	..	..	..	..	1	2	..	..	2	3	1	..	6	4	1
22	..	..	..	..	..	1	..	..	..	..	..	..	..	2	..	..	2	5	..	1	6	3	..
23	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	18	..	1	1	..

TABLE II. PROBLEMS RANKED ACCORDING TO REAL DIFFICULTY AND TEACHERS' ESTIMATES

Problem	*Rank (real difficulty)*	*Average rank of teachers' estimates*
1	3	2
2	6	4
3	10	1
4	18	9
5	2	6
6	20	18
7	5	11
8	4	5
9	11	7
10	9	8
11	1	3
12	16.5	13
13	19	14
14	15	17
15	8	12
16	13	10
17	14	19
18	22	20
19	23	23
20	12	15.5
21	16.5	22
22	21	21
23	7	15.5

Table I shows that four teachers judged Problem 1 to be the easiest, five judged it to be second in difficulty, two judged it to be third in difficulty and so on. The remarkable thing about this table is the lack of agreement in the judgments of the teachers. The problems are not unusual. The list or a portion would make a typical examination. It should be clear that if a teacher giving such a list attempts to assign values to the several questions, the values thus assigned are likely to be inaccurate. It is doubtful, except in extreme cases such as Problem 19, whether the values will be more accurate than if the questions are considered to be equal in value.

The problems were given to the pupils in the fifth, sixth, seventh, and eighth grades in one school. The total number of pupils was about 1500. By doing this it was possible to determine what problem actually was easiest, and the order of difficulty for the entire list. The average rank of each problem as determined by the pupils' scores, and the rank as determined by the average of teachers' opinions, are given in Table II. This table presents some interesting facts. Twelve of the twenty teachers agreed that Problem 3 was easiest and no one ranked it above seventh in difficulty. Its real rank was found to be tenth. Similar discrepancies can be pointed out for other problems, although in the case of certain problems the average of the teachers' estimates approximates the real rank. Thus even the average judgment of twenty teachers on the relative difficulty of problems is not reliable and the judgment of a single teacher is much less reliable.

(2) *Rate of doing work neglected.* In the second place, it is customary in giving an examination to allow sufficient time for all pupils to answer all of the questions, or if this is not done, the papers are graded on the basis of what each pupil has done. This manner of giving an examination fails to take into account the rate at which a pupil is able to answer the questions. Only the quality of the answers is considered, and the pupil who answers the questions with difficulty, and who barely finishes in the time allowed, receives exactly the same "grade" as the more capable pupil who is able to answer the questions easily and who finishes in one half or one third of the time, providing the two sets of answers are equivalent. It is clear that when this is done, the "grade" or mark which the pupil receives is not a true measure of his ability, because the rate at which he is able to do work is a "dimension" of his ability as well as the quality of what he does. In certain cases the rate may be a relatively unimportant dimension. Neglecting it in measuring the ability of a pupil is much like neglecting the width in measuring a rectangle to determine its area.

Some may insist that it is unfair to the slow-working pupil not to allow sufficient time for him to answer all of the questions. However this may be, it certainly is unjust to the more capable pupil to deprive him of the opportunity to demonstrate what he is able to do. This is exactly the case when the work asked of him is sufficient to keep him employed only a half or a third of the period allowed for the examination. This practice of ignoring the rate of working probably tends to cause desultory and careless school work.

Investigation has shown that rapid work and a high degree of quality or accuracy are not incompatible in arithmetic. The same statement can be made with reference to reading. Investigation has indicated that a considerable per cent of pupils can be made more accurate in arithmetic by forcing them to work more rapidly. It has also been shown that about three pupils out of four make progress in rate of work and accuracy at the same time. In view of these facts, it appears that good instruction requires that the teacher give attention to the rate of doing work as well as to the quality of the work done. The rate at which a pupil is able to do work of a given quality is as much a factor of his ability as is the quality of the work which he does.

The rate at which a pupil works can be measured very easily. It is simply necessary to secure a record of the time which he spends in answering the set of questions. When an examination is given to a group, it is rather inconvenient to secure a record of the time which each pupil spends upon the examination. However, one can secure just as true a record of the rate at which each pupil works by making the examination long enough so that no pupil finishes in the time allowed. For each pupil the number of minutes, divided by the number of units of work which he did, will give his rate of working per unit.

Summary. In the preceding pages we have shown, first, that questions differ in difficulty and that teachers cannot judge their relative difficulty with reliability; and, second, that the rate of doing work, which is in many cases an important "dimension" of ability, is commonly neglected in giving examinations. The first of these conditions contributes to the inaccuracy of marking examination papers. The second means that the examination paper frequently is not a true record of the pupil's ability. This happens when the pupil finishes before the end of the period allowed. There are two other points which should be mentioned in this connection. Marks placed upon examination papers do not have a definite meaning because a wide range of topics is included within a single examination and because no reliable standards exist.

Wide range of topics included within an examination makes the "grade" have an indefinite meaning. Examinations are usually made up of questions from a number of different fields within a subject. Take, for example, the following examination in arithmetic which was given to a sixth-grade class:

1. Write in Roman system: 49, 79, 94, 96, 146.
2. If 11 A. of land are worth $1485, what is one acre worth?
3. If a desk is 4 2/3 ft. long and 3 5/12 ft. wide, what is the perimeter?
4. How much must you add to 26 7/8 in. to make a yard?
5. A man has to travel 117 mi. After going 5/9 of the distance, how many miles has he still to travel?
6. The perimeter of a square is 851 in. What is the length of one side?
7. Of 152 chickens a hawk captured 12 1/2%. How many were captured? How many were left?
8. A man saves $675.20 a yr., which is 32% of his income. How much is his income?
9. At $1.38 a yd., what will 37 yds. of carpet cost?
10. At $65.50 an acre, what must a man pay for 25.4 acres of land?

Question 1 calls for a knowledge of Roman numerals; Question 2 asks the pupil to find the cost of a unit when the cost of the whole is given; Questions 3, 4, and 6 deal with mensuration; Question 5 calls for the finding of a fractional part of the whole; Questions 7 and 8 are problems in buying. Thus we find six different topics included within an examination of ten questions.

Suppose a pupil receives a "grade" of 80 on this examination. Even if 80 is an accurate measure of what the pupil is able to do on this examination, it cannot have a definite meaning. It does not tell us whether the pupil lacks ability in the field of Roman numerals, or in the field of percentage, or in some other of the fields included in this examination. In order that the total score made on an examination may be a definite measure of a pupil's ability, the questions which compose it must be drawn from a single field, or at most from a small group of closely related fields. If this is not done, the scores for each question must be kept separate in order to have a definite meaning.

The situation is much the same as if the length, width, height, seating capacity, number of windows, and the number of doors of a room were added together to form a measure of the room. If we assume that each of these characteristics of the room was measured with a high degree of accuracy, the total of the numbers expressing the measures gives us only very general information about the room. If the total is large, we know that the room is probably large; if the total is small, we know that it is small. But under no circumstances can we be certain that the room has any windows or doors, that it contains any seats, or that its dimensions are well proportioned. In order that we may have definite information about the room, it is necessary that the measures of the several characteristics be kept separate.

No standards for interpreting measures. The fact that a seventh-grade pupil solves correctly eight problems out of seventeen or spells correctly twenty-one words out of twenty-five has a meaning only by comparison with the *standard* for these examinations. By standard we mean the number of problems which a pupil of a given grade, in this case the seventh, should do correctly when given this examination. If the standard is twelve problems, this pupil is below seventh-grade standard in ability and has not done satisfactory work. On the other hand, if the standard is six problems, this pupil is above standard and possesses superior ability. Without a standard a teacher cannot know what a measure means.

The above statement may not appear to be true at first thought. Standards have not been determined for the examinations which a teacher gives, but he "guesses" what the standard should be when the questions are made out and the examination is judged by the teacher to be "fair" for the pupils of that grade or one "they should be able to pass." We have just seen how unreliable teachers' judgments are with reference to the difficulty of problems in arithmetic. Their "guesses" with reference to standards appear to possess about the same degree of reliability.

Accurate measurements of the abilities of pupils may be made by using standardized tests. The preceding pages were written to make clear that our present measurements of the abilities of pupils were inaccurate and hence unsatisfactory. Since the measurement of results is very necessary to both the teacher and the supervisor, there is a need for instruments with which accurate measurements can be made. Standardized tests are such instruments and in the following chapters certain ones will be described and directions given for their most effective use by the teacher.

Standardized tests have been scientifically devised. The questions or exercises which make up the tests have been carefully selected and evaluated. Directions have been provided so that different teachers will assign the same mark to the same paper. The rate of work is measured where it is an important "dimension" and the tests have been standardized. Generally a standardized test is limited to a single topic or to a small group of topics so that a pupil's score has a definite meaning. These features eliminate the defects in ordinary examinations which have been discussed in this chapter and hence constitute reasons for the use of standardized tests.

Other advantages in using standardized tests. Standardized tests are helpful in another way to the teacher, particularly the rural teacher who must work isolated for the most part from other teachers. The standards of such tests are definite objective aims stated in a way that both teacher and pupil can understand. The value of a definite standard can hardly be overestimated. As we shall show later it furnishes a strong motive. It also guides one's efforts. It makes possible economy of time by limiting training. The use of standardized tests directs attention to the results which are to be attained. Too often attention has been focused upon the method being used rather than upon the results. A third advantage is due to the fact that the patrons of the school are interested in definite statements of results, particularly when those results can be compared with recognized standards. Many objections to a teacher or a school have been answered by the accurate measurement of results. The writer has heard one superintendent state that standardized tests would be worth using if they did nothing more than stop the mouths of those who are accustomed to complain about what the public school is doing.

QUESTIONS AND TOPICS FOR STUDY

1. What evidence do we have for showing that "final grades" are inaccurate measures of the abilities of pupils?
2. How do we know that the marking of examination papers is inaccurate?
3. What factors contribute to this inaccuracy? Can you think of any not mentioned in this chapter?
4. Have you ever felt that examination marks were inaccurate? Why?
5. Ask several teachers to "grade" the same set of papers and compare the "grades" given to each paper?
6. What is meant by saying that a "grade" has an indefinite meaning?
7. What is a standard? Why are standards needed?
8. What are the advantages of using standardized tests?
9. The unreliability of individual judgment may be shown by having a group of persons guess the length of a stick when it is held as much as ten feet away from them.
10. What is meant by saying that the rate of doing work is a "dimension" of a pupil's ability? Why is it important to measure it?

CHAPTER II

THE MEASUREMENT OF ABILITY IN READING[1]

THERE are two types of reading, silent reading and oral reading. In reading silently one is concerned primarily with understanding the printed page. In oral reading the point of emphasis is the communication of the meaning by means of oral expression. Both kinds of reading are taught in the school, and the first step in the measurement of ability in reading is to recognize the existence of the two types of reading ability. We shall consider their measurement in two separate sections.

I. SILENT READING

1. Monroe's Standardized Silent Reading Tests

Ability to read silently is measured by having the pupil read a selection and then give evidence of the degree of his understanding or comprehension of the material read. One method of securing this evidence is to require the pupil to answer one or more questions based upon what he has read. This plan may be illustrated by the following paragraph and question. Answering the question requires that the pupil comprehend the principal idea of the exercise.

Not far from Greensburg is a little valley, among the high hills. A small brook glides through it, with just murmur enough to lull one to repose; and the occasional whistle of a quail, or tapping of a woodpecker, is almost the only sound that ever breaks in upon the uniform tranquillity.

---

[1] The reader should have a copy of each of the standardized tests described in this and the following chapters. In several instances it will be almost impossible to understand the discussion without a copy of the test at hand. See the Appendix for directions for securing a sample package and for purchasing any of these tests for class use.

What kind of a picture do you get from reading the above paragraph?

disorder activity noise calmness confusion

This exercise is expressed in such a way that the pupil will have no difficulty in expressing his answer if he knows what it is, and also his answer will be either right or wrong. There can be no difference of opinion in marking the exercise. A series of tests, known as Monroe's Standardized Silent Reading Tests, consists of a number of such exercises. The exercises were taken from school readers and other books that children read which insures that they present typical reading situations. The amount of credit to be given a pupil for doing each exercise correctly has been scientifically determined and is called the *comprehension value.* The sum of the comprehension values of the exercises done correctly in five minutes makes the pupil's *comprehension score.* This score is the measure of his ability to comprehend or understand the exercises of the test.

The pupil's rate of reading is important as well as the degree of his understanding. For this reason each exercise has a rate value, and a pupil's *rate score* is the sum of the rate values of the exercises which he tries in five minutes regardless of whether he does them correctly or not. This value has been so chosen that it represents the number of words which the pupil reads per minute. A pupil's *rate score* is the measure of his rate of reading.

Test I of the series is for Grades III, IV, and V, Test II for Grades VI, VII, and VIII, and Test III for Grades IX to XII. There are three forms of Tests I and II which are equivalent in difficulty, so that when it is desired to measure the ability of the pupils a second or third time it is not necessary to use the same exercises. A few exercises of Test II are reproduced to illustrate more fully this type of silent reading test.

Rate Value 7

No. 2

At evening when I go to bed
I see the stars shine overhead;
They are the little daisies white
That dot the meadow of the night.

What are the little white daisies of the night?

Comprehension Value 2.1

Rate Value 8

No. 4

They rested and talked. Their talk was all about their flocks, a dull theme to the world, yet a theme which was all the world to them.

What do you suppose was the occupation of these men?

carpenter doctor merchant
shepherd blacksmith

Comprehension Value 2.7

Rate Value 11

No. 7

He was a wicked ruler who, with his still more wicked sons, oppressed and wronged the people in many ways.

If the people would be sorry when the ruler and his sons died, draw a line under the word ruler; if they would be glad, cross out the word ruler.

ruler

Comprehension Value 2.0

Rate Value 17

No. 10

It was cold, bleak, biting weather; foggy withal; and he could hear the people in the court outside go wheezing up and down, beating their hands upon their breasts and stamping their feet upon the pavement-stones to warm them.

The author has attempted to give you a picture in this paragraph. After reading the paragraph, if you think it is a picture of comfort and pleasantness, draw a line under the word hear; if of cheerlessness and dreariness, draw a line under bleak.

hear wind bleak cold

Comprehension Value 4.3

Directions for using the tests. Detailed directions for giving these tests are printed on the first page of the test paper and hence it is not necessary to reproduce them here. However, there are four general rules which should be followed in the giving of all standardized tests: (1) Follow the printed directions carefully. Do no more or no less than the directions specify. Do not try to improve upon the directions. Comparisons of the scores of your pupils with the scores of other classes and with the standard scores will not be valid if the printed directions are not followed, because these scores were obtained according to these conditions. (2) Be careful to allow exactly the number of minutes specified — five minutes. Use a watch with a second-hand or a stop-watch if one is available. (3) The examiner should exercise care not to excite or frighten the pupils by his manner of giving the tests. He should not be in a hurry. He should not be cross. He should remember that reliable measurements of the abilities of the pupils will not be obtained unless the pupils work naturally. (4) Study the directions for the tests until you are familiar with them. It is wise to go through the directions at least once imagining that you have the class before you. Your failure to be familiar with the directions may affect the scores of your pupils.

Giving the tests in rural schools. These silent reading tests may be given to a group of pupils belonging to several different grades as easily as to a group belonging to a single grade. It is only necessary to see that each pupil is provided with the test which is designed for his grade. The time allowance is the same for all grades. In a rural school it will be most convenient to test all of the pupils above the second grade at one time. In recording the scores it will, of course, be necessary to record the scores for the different grades separately.

When the tests should be given. These silent reading tests are not teaching devices. They are instruments for measuring the ability of pupils to read silently. They should be given at the beginning of the school year so that the teacher may know his pupils better. If they are not used at the beginning of the year, they may be given at any time, preferably as early as convenient. The tests should be repeated at the end of the year so that the teacher may know how much his pupils have increased their ability to read silently. When the tests are given a second time a different form should be used. If it is desired, the tests may be given a third time at the middle of the year, but they should not be given more than three times a year.

Scoring the test papers. The correct answer for each exercise is given on the back of the class record sheet which is always furnished with the tests. It is most satisfactory for the teacher to mark the papers, but if the teacher feels that he cannot take the time for it he may read the answers and have the pupils mark their own papers, or better, have them exchange papers. In any case the teacher should examine enough of the papers to make certain that they have been marked correctly. The question has been asked, "Should the pupils be required to give their answers in the form of complete sentences?" This is not required. The author does not believe that it is wise to insist upon this form.

Good arrangement of scores. The significance of a group of facts, such as the scores made by a class upon a test, may be made more evident by certain methods of arranging them. Take, for example, the comprehension scores which were made by a sixth-grade class of thirty-five pupils when given a certain silent reading test. When these scores are presented in the manner of Table III, the array tends to confuse. One must scan the entire array to learn that the lowest score is 4.2, or that the highest score is 30.1. One cannot easily learn that pupil BB, who made a score of 14.9, stands eighth from the poorest in the group. If now the scores are simply rearranged in order of magnitude, as shown in Table IV, their significance is much more easily grasped.

Table III. Showing a Poor Arrangement of Scores

Pupil	*Score*	*Pupil*	*Score*	*Pupil*	*Score*
A	27.3	M	10.0	Y	16.0
B	19.2	N	16.3	Z	19.1
C	26.2	O	21.1	AA	15.4
D	22.5	P	25.6	BB	14.9
E	15.4	Q	21.1	CC	16.4
F	18.3	R	15.9	DD	14.1
G	28.4	S	16.1	EE	4.2
H	17.4	T	5.9	FF	20.0
I	25.1	U	30.1	GG	24.1
J	15.7	V	22.3	HH	26.3
K	11.8	W	13.1	II	25.8
L	21.6	X	12.8		

Table IV. Showing the Same Scores rearranged in a Better Order

Pupil	*Score*	*Pupil*	*Score*	*Pupil*	*Score*
EE	4.2	Y	16.0	V	22.3
T	5.9	S	16.1	D	22.5
M	10.0	N	16.3	GG	24.1
K	11.8	CC	16.4	I	25.1
X	12.8	H	17.4	P	25.6
W	13.1	F	18.3	II	25.8
DD	14.1	Z	19.1	C	26.2
BB	14.9	B	19.2	HH	26.3
AA	15.4	FF	20.0	A	27.3
E	15.4	O	21.1	G	28.4
J	15.7	Q	21.1	U	30.1
R	15.9	L	21.6		

Recording the scores. For securing a good arrangement of the scores obtained by using Monroe's Standardized Silent Reading Tests the class record sheet shown on page 28 is used. It will be noted that the scores are arranged in order of magnitude by groups. This kind of an arrangement of scores is called a *distribution*. For comprehension, all of the scores from 3.0 to 3.9 are grouped together. The dif-

Rate Score		Comprehension Score	
Interval	Number of Pupils	Interval	Number of Pupils
Above 160	....		....
141 to 150	....		....
151 to 160	....		....
131 to 140	....		....
121 to 130	....	80 & above	....
116 to 120	....	70 to 79.9	....
111 to 115	....	60 to 69.9	....
106 to 110	....	50 to 59.9	....
101 to 105	....	45 to 49.9	....
96 to 100	....	40 to 44.9	....
91 to 95	....	35 to 39.9	....
86 to 90	....	30 to 34.9	....
81 to 85	....	27 to 29.9	....
76 to 80	..71......	24 to 26.9	....
71 to 75	....	21 to 23.9	....
66 to 70	....	18 to 20.9	..19.45......
61 to 65	....	15 to 17.9	....
56 to 60	....	13 to 14.9	....
51 to 55	....	11 to 12.9	....
46 to 50	....	9 to 10.9	....
41 to 45	....	7 to 8.9	....
36 to 40	....	5 to 6.9	....
31 to 35	....	4 to 4.9	....
26 to 30	....	3 to 3.9	....
21 to 25	....	2 to 2.9	....
16 to 20	....	1 to 1.9	....
Below 15	....	0 to .9	....
Total		Total	
Median		Median	

FIG. 4. SHOWING FORM USED IN RECORDING THE SCORES OBTAINED BY USING MONROE'S STANDARDIZED SILENT READING TESTS.

ference between 3.0 and 3.9 (more exactly 3.9999−) or 1, is called the *width of the interval.* On this record sheet all of the intervals do not have the same width. For example, the interval from 24.0 to 26.9 has a width of 3. Detailed directions for recording the scores are printed on the class record sheet and need not be repeated here.

The central tendency of a distribution. Ordinarily the scores of a class will be distributed over several intervals of the class record sheet. If one wishes to compare the standing of the class as a whole with the standard, it is necessary to obtain a *central tendency of the distribution.* The central tendency which is best known by teachers is the average, but if one or two pupils make very low scores they will bring the average down. For this reason the *median* is used instead of the average. The *median* is the value of the middle score of the distribution. The median score for comprehension is found by arranging the test papers according to the size of the comprehension scores. When the test papers are arranged in order, the score on the middle paper is the median score. For example, if there are thirty-five papers in the pile, the score on the eighteenth paper is the median score. If there are thirty-six papers, the median score is halfway between the score on the eighteenth paper and the score on the nineteenth paper. The median score for rate is found the same way. The median scores are called the *class scores.*

Summary. We have described Monroe's Standardized Silent Reading Tests, the directions for giving them, for recording the scores and for finding the class scores. In the next chapter we shall take up the meaning or interpretation of the scores and what a teacher should do to correct the conditions which the tests reveal.

2. Courtis's Silent Reading Test No. 2

Description of the test. This test, which is to be used in Grades 2 to 6 inclusive, is designed to measure "the ability to read silently and understand a simple story and simple questions about the story." It consists of a connected story of the kind that children enjoy reading. The first two paragraphs of Part I of the test are reproduced on page 30 to show the type of story and its arrangement.

The pupils are directed: "Read *silently*, and only as fast as you can get the meaning; for when you have finished you will be asked to answer questions about what you have read. You will be marked for both, how much you read and how

The Kitten Who Played May-Queen

5 When the spring sun was warm and the
13 spring flowers had begun to bloom, our children
21 had a May-party out on the lawn. Every little
30 boy or girl who lived nearby was invited. Betty
39 wrote the invitations and Bobby carried them to
47 the children.

49 When the day of the party came, Daddy planted
58 a May-pole and Mother tied it with gay-colored
66 ribbons. There were to be games and dances
74 on the grass and a delicious supper, with
82 a basket full of flowers for every child.

well you understand it, but it is better to get the meaning of the story than to read too fast."

The pupil reads silently for three minutes, but at the end of each half-minute a signal is given, and he is to mark the last word he has just read and to keep on reading. At the end of three minutes the pupil turns to Part II. This consists of the same story with questions based upon it. The first two paragraphs and the questions upon them are reproduced.

When the spring sun was warm and the spring flowers had begun to bloom, our children had a May-party out on the lawn. Every little boy or girl who lived nearby was invited. Betty wrote the invitations and Bobby carried them to the children.

A. Did the children have a May-party? ....______
B. Was it Bobby who wrote the invitations? ....______
C. Was the party held in the house? ....______
D. Were only girls invited to the party? ....______
E. Had the spring flowers begun to bloom? ....______

When the day of the party came, Daddy planted a May-pole and Mother tied it with gay-colored ribbons. There were to be games and dances on the grass and a delicious supper, with a basket full of flowers for every child.

1. Were the children to have anything to eat? ....______
2. Were they going to play on the grass? ....______
3. Were they going into the house to dance? ....______
4. Were the baskets to be full of flowers? ....______
5. Was it Daddy who tied the ribbons to the pole? ....______

The measures of the pupil's ability to read silently and answer questions. The pupils are given five minutes to answer as many of the questions as they can. The measure of their understanding of the story is expressed in terms of the *number of questions answered* and the *index of comprehension*. This index is found as follows: "Subtract the wrong answers from the right answers. (If there are more wrong than right, find the difference and give it a negative sign.) Divide the difference by the number of right answers, carrying the result to three places and keeping two." This calculation is made from a table so that the labor involved is reduced to a minimum.

In addition to these two scores, the *number of words read per minute* is obtained from the first part of the test. This is a measure of the pupil's rate of reading.

Directions for using Courtis's Silent Reading Test No. 2. Detailed directions for giving the test have been prepared by the author of the test. These do not need to be repeated here, but the general directions mentioned on page 25 apply to these tests as well. Courtis says:

> The instructions given in this folder must be followed exactly if the results secured are to be compared with results from the same tests in other schools. The examiner needs to prepare himself for his work by careful study and practice.
>
> The important conditions to be controlled are *timing*, *instructions*, and *manner*. Keep *exact* times. *Say no more* to the children about the tests than is provided for in the instructions below. Give the instructions energetically, but easily and *pleasantly*. Do not hurry, do not get excited, do not be cross with the children. The children will make the best scores if they enjoy the tests and work naturally.

Directions for using Courtis's Silent Reading Test No. 2 in rural schools. Since the same test is given to all pupils in Grades 2 to 6, the procedure for giving the test is always the same and therefore it may be used as easily in rural schools as in graded schools. The only difference comes in recording the scores. Then the teacher must first be careful to have the test papers grouped correctly by classes or grades.

Scoring the test papers. An answer card and detailed directions are furnished with the tests. The most accurate results are obtained when the test papers are scored by the teacher, but Courtis says: "A teacher might better give her time to studying the results than waste it on scoring." His plan is to have the scoring done by the pupils in the fifth and sixth grades and by the seventh- and eighth-grade pupils in the case of pupils below the fifth grade. The scores of a pupil are recorded on an individual record card which is more convenient to handle in recording the scores of the class on the class record sheet.

Recording the scores. The three scores, rate of reading (words per minute), number of questions, and the index of comprehension are recorded on the forms marked Table 1 and Table 3 in Fig. 5. The form of Table 1 is similar to that used for recording the scores of Monroe's Standardized Silent Reading Tests. However, it should be remembered that in Table 1, 0 means 0 to 19, 20 means 20 to 39, and so on. Table 3 represents a different type of form for recording scores. The test papers, or in this case the individual cards containing the record of the pupils' scores, are first sorted into piles according to the number of questions answered, putting into the first pile all the cards having a score of 0 to 4 questions answered, into the second pile all those cards having a score of 5 to 9 questions answered, and similarly for the other intervals. Then each of these piles is sorted according to the "index of comprehension." Suppose there are seven cards in the pile of forty-five to forty-nine questions answered and that these have the following "indices of comprehension": −15, 20, 47, 58, 73, 79, 80. The entries for these scores would be made on the "45 line." To record the −15 score, a 1 is placed in the "Less than −5" column; to record the 20 score a 1 is placed in the "6–39" column; to record the 47 and 58 scores a 2 is placed in the "40–69" column; and so on.

Distributions and central tendencies. The column and line marked "Total" in Table 3 are filled in by finding the sum of the scores recorded in the respective lines or columns. These two sets of sums form the distributions for "number of questions answered" and "index of comprehension."

Table 1

Rate of Reading

Score in words per minute	Number of children making each score
Over	
400	
380	
360	
340	
320	
300	
380	
360	
340	
320	
300	
280	
260	
240	
220	
200	
180	
160	
140	
120	
100	
80	
60	
40	
20	
0	
Total	
Median	

Table 3

Index of Comprehension

Diagnosis		Guesswork			Comprehension poor; additional training needed					Comprehension satisfactory		
Questions Answered	Total	Less than −5	5 to +5	6–39	40–69	70–79	80–84	85–89	90–94	95–99	100	
70												
65												
60												
55												
50												
45												
40												
35												
30												
25												
20												
15												
10												
5												
0												
Total												

Rate of work satisfactory.

Diagnosis: Rate of reading too slow. Additional training needed.

Median Number of Last Question Answered........

Median Index of Comprehension........

FIG. 5. SHOWING FORMS USED IN RECORDING THE SCORES OBTAINED BY USING COURTIS'S SILENT READING TEST NO. 2.

The numbers in the column headed "Number of children making each score" in "Table 1" form the distribution for the rate of reading. The central tendency of the distribution which is used to describe the achievement of the class as a whole is the median the same as was used for Monroe's Standardized Silent Reading Tests. It may be found in the same way, although the directions given by Courtis differ slightly.

Summary. In the preceding pages we have described Courtis's Silent Reading Test No. 2 and the general directions for giving it and for handling the scores. Detailed directions are given in the folders, B and D, which should be secured when the test is purchased for classroom use.

Comparison of the two silent reading tests. In comparing the two silent reading tests which we have just described, it must be remembered that they probably do not measure the same type of silent reading ability. We cannot, therefore, compare them as we might two things of the same kind, as we might two makes of the same type of automobile. The situation is much like comparing a pleasure car with a truck. However, certain points may be noted. Courtis's test is to be used in Grades 2 to 6, while Monroe's tests are to be used in Grades 3 to 8. Thus a second-grade teacher would choose Courtis's test. Some claim that it is more satisfactory for the third and fourth grades, but that Monroe's test should be used beginning with the sixth grade. Of the two tests, Monroe's takes less time and is simpler to understand and use. For the teacher who is inexperienced in the use of standardized tests this is an important consideration. However, it is frequently profitable to give both tests. One will supplement the other.

3. Thorndike's Visual Vocabulary Scale

Need for a vocabulary test. When a pupil makes a silent reading score which is below standard, it may be due to any one of several causes or to a combination of causes. One important cause and one which frequently occurs is the pupil's failure to comprehend the meaning of the words. Understanding the meaning of a paragraph requires the ability to comprehend the meaning of the individual words. Hence a vocabulary test furnishes a means for obtaining information which may reveal the cause of the pupil's low score in silent reading.

Description of the test. This test can be most satisfactorily described by reproducing a portion of it. Each pupil is given a test paper on which are printed the following directions and lists of words.

THORNDIKE READING SCALE B. WORD KNOWLEDGE OR VISUAL VOCABULARY — SERIES X

Write the letter W under every word that means something about **war** or **fighting.**

Write the letter B under every word that means something about **business** or **money.**

Write the letters CHU under every word that means something about **church** or **religion.**

Write the letter R under every word like **father** or **wife** that means something about **relatives** or the **family.**

Write the letters COL under every word that means a **color.**

Write the letter T under every word like **now** or **then** that means something to do with **time.**

Write the letter D under every word like **here** or **north** that means something about **distance** or **direction** or **location.**

Write the letter N under every word like **ten** or **much** that means something about **number** or **quantity.**

4.0 camp, flag, west, mother, two, general, green

troops, south, fort

4.5 gray, cousin, pink, uncle, yellow, hour, pay, aunt, early, commander

5.0 marriage, defeat, many, afternoon, guard, buy, captive, military, relation, late

6.0 hymn, defend, across, merchant, noon, forty, conquer, dagger, profit, Tuesday

There are eight other lists which gradually increase in difficulty. The scale values of the several lines printed in the left margin were determined by having several thousand children undertake to indicate the meaning of each word. The greater the per cent of children who could not indicate correctly the meaning of the word, the higher the value attached to the word. The pupil's score is the value of the most difficult list in which he marks correctly eight out of the ten words. Likewise the class score is the value of the list for which the class averages 80 per cent (8 out of 10) of correct meanings.

Giving the test. Detailed directions for giving this test are printed on the cover of the test.[1] It is, therefore, necessary only to remind the reader that the general directions on pages 25 and 32, with reference to following directions and the manner of presenting the tests to the pupils, apply.

Recording the scores. The class record sheet devised by Thorndike called for a detailed record for each pupil. This was cumbersome to use and required more labor than appeared to be justified for obtaining the class score. A simpler class record sheet has been designed by the writer. This is reproduced in Fig. 6. Directions for recording the scores on

---

[1] This is not true of the form of the test which is secured from the Bureau of Publications of Teachers College, but is true of the form distributed by the Bureau of Educational Measurements and Standards, Emporia, Kansas.

Line value	Number of words wrongly marked or omitted	Total	Per cent
4.0	....	....	....
4.5	....	....	....
5.0	....	....	....
6.0	....	....	....
6.5	....	....	....
7.0	....	....	....
7.5	....	....	....
8.0	....	....	....
8.5	....	....	....
9.0	....	....	....
9.5	....	....	....
10.0	....	....	....

Class Score..............

Fig. 6. Record Sheet for recording Scores obtained by using Thorndike's Visual Vocabulary Scale.

this sheet and for calculating the class score are printed on the back of it and thus need not be reproduced here. The calculation of the class score may appear to be difficult to understand, but if the directions are studied carefully and followed a step at a time, one should soon learn how to do it.

II. ORAL READING

1. Gray's Oral Reading Test

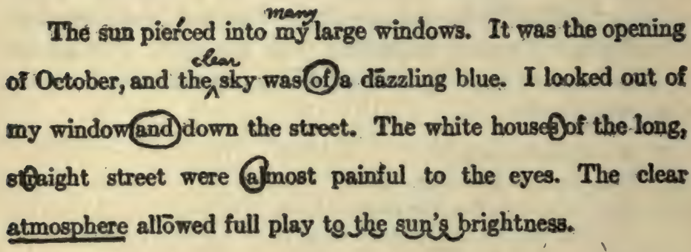
Description of the test. For measuring the ability of pupils to read orally, Gray has devised a test consisting of a series of paragraphs arranged in order of gradually increasing difficulty in oral reading. The test is designed to be used in all grades beginning with the first. The nature of the test can best be illustrated by reproducing a few of the paragraphs:

1. A boy had a dog.
The dog ran into the woods.
The boy ran after the dog.
He wanted the dog to go home.
But the dog would not go home.
The little boy said, "I cannot go home without my dog."
Then the boy began to cry.

6. It was one of those wonderful evenings such as are found only in this magnificent region. The sun had sunk behind the mountains, but it was still light. The pretty twilight glow embraced a third of the sky, and against its brilliancy stood the dull white masses of the mountains in evident contrast.

11. The hypotheses concerning physical phenomena formulated by the early philosophers proved to be inconsistent and in general not universally applicable. Before relatively accurate principles could be established, physicists, mathematicians, and statisticians had to combine forces and work arduously.

Giving the test. In giving the test the pupils are taken one at a time and asked to read beginning with the first paragraph. As the pupil reads, the teacher records on another copy of the test, two sets of facts: the number of seconds the pupil takes to read each paragraph and the errors which are made. To obtain the number of seconds the teacher must have a watch with a second-hand or, better, a stop-watch. Six types of errors are recorded: (1) complete mispronunciation of a word so as to indicate that the pupil has no control over it; (2) partial mispronunciation; (3) omissions; (4) substitutions; (5) insertions; and (6) repetitions. The method of marking these errors on the test paper is illustrated in the following quotation from the class record sheet:



> **The sun pierced into my large windows. It was the opening of October, and the sky was of a dazzling blue. I looked out of my window and down the street. The white houses of the long, straight street were almost painful to the eyes. The clear atmosphere allowed full play to the sun's brightness.**
>
> If a word is wholly mispronounced, underline it as in the case of "atmosphere." If a portion of a word is mispronounced, mark appropriately as indicated above: "pierced" pronounced in two syllables, sounding long *a* in "dazzling," omitting the *s* in "houses" or the *al* from "almost," or the *r* in "straight." Omitted words are marked as in the case of "of" and "and"; substitutions as in the case of "many" for "my"; insertions as in the case of "clear" and repetition as in the case of "to the sun's." Two or more words should be repeated to count as a repetition.

To give the test satisfactorily requires practice in detecting the errors and in recording them. The teacher should have some one read the test, intentionally making errors, so that he may become skillful in giving it.

The pupil's score. The pupil's score depends upon both the number of seconds he takes to read the different paragraphs and the number of errors which he makes. For example, certain credit is given a second-grade child for reading paragraph 1 in forty seconds with less than five errors, and additional credit is given the same child for reading the same paragraph in thirty seconds with less than five errors, or in forty seconds with less than four errors. Still different credit is given to third-grade children for each of the above achievements with paragraph 1. When the combination of length of time and number of errors exceeds a certain prescribed maximum, no credit is allowed. The score of any child is ascertained by adding together all the credits which he has earned on the several paragraphs. This process becomes much more simple than it sounds here when explicit directions and the detailed data for each child and for tabulating results are at hand.

Silent reading versus oral reading. Notwithstanding the fact that oral reading has received much greater emphasis in our schools than silent reading, the latter is far more important. Silent reading is required in practically all of the other school subjects. Also the pupil will read silently much more frequently than orally after he leaves school. However, at first oral reading is a means for teaching silent reading. Hence in the primary grades it is worth while to measure the ability of pupils to read orally.

Summary. In this chapter we have described two silent reading tests, one vocabulary test and one test for oral reading. The detailed instructions for using these tests have not been reproduced since they always are furnished with the tests. We have given only those general directions which were considered necessary for understanding the tests and the scores obtained by using them. Certain words which are used in discussing educational tests have been introduced. The reader should study the meaning of these words carefully because they will be used frequently in the following chapters. The most important of these words are: *score, distribution of scores, central tendency, median, class scores.* In the next chapter we shall consider the meaning or interpretation of scores and how to correct the defects revealed by the tests.

QUESTIONS AND TOPICS FOR STUDY

1. What is silent reading? What is oral reading? Which is the more important? Why?
2. What are the essential features of Monroe's Standardized Silent Reading Tests?
3. What are the essential features of Courtis's Silent Reading Test No. 2?
4. What is a distribution of scores?
5. What is the median score?
6. How is the median score found in using Monroe's Standardized Silent Reading Tests?
7. Why should a vocabulary test be used?
8. Have you been satisfied with your marking of pupils in reading? How have you tested pupils in reading? Do you think you have done it as well as you could by using the tests described in this chapter?
9. Have you been placing too much emphasis upon oral reading? How could you find out?

CHAPTER III

THE MEANING OF SCORES AND CORRECTING DEFECTS IN READING

I. MONROE'S STANDARDIZED SILENT READING TESTS

Standards necessary to give scores meaning. Seven seventh-grade pupils made the scores in Table V when given Monroe's Standardized Silent Reading Test in April. Although it is obvious that certain of these scores are larger than others none of them mean very much until we know what scores a seventh-grade pupil should make. That is, standard scores are necessary for interpreting the scores of pupils or classes.

TABLE V. SHOWING SCORES OF SEVEN SEVENTH-GRADE PUPILS ON MONROE'S STANDARDIZED SILENT READING TEST.

Pupil	*Comprehension score*	*Rate score*
A. N.	35.0	146
E. A.	27.6	98
C. S.	23.1	146
H. H.	22.8	146
R. H.	17.8	85
F. S.	14.5	54
E. P.	11.8	98

These tests have been given to several thousand pupils in each grade and the resulting scores tabulated as the scores of a class are recorded. (See form on page 28.) From these distributions it is a simple matter to calculate the scores which the "average" or typical pupil in each of the grades makes. These scores are *standard scores.* Table VI gives the standard May scores for these tests; that is, the scores which the "average" pupils completing the respective grades make. When the tests are given at the beginning of the school year, one would use the standards of the grade below for judging the pupils' scores. In case the tests are given at some time during the school year, say at the end of the fourth month, approximate standard scores for this date can be calculated from the facts of Table VI.

TABLE VI. STANDARD MAY SCORES FOR MONROE'S STANDARDIZED SILENT READING TESTS

	Test I			*Test II*		
Grade	III	IV	V	VI	VII	VIII
Comprehension	9.0	14.5	21.0	21.0	24.0	27.5
Rate	60	80	93	92	102	108

The seventh-grade scores in Table V were obtained by giving the tests April 15 which is about the end of the eighth month of school, but since the difference between the sixth- and seventh-grade standards is not large, we can use the May standards without introducing an appreciable error. Pupil A. N. (scores 35.0, 146) is distinctly above standard in silent reading ability as shown by this test. Pupil E. A. (scores 27.6, 98) is above in comprehension, but slightly below in rate of reading. Pupils C. S. and H. H. (scores 23.1, 146; 22.8, 146) are approximately standard in comprehension and read very much faster than the standard rate. The other three pupils are below standard in both comprehension and rate. Pupil E. P. (scores 11.8, 98) is close to the standard in rate, but his comprehension score is less than half of the standard. Pupil F. S. (scores 14.5, 54) reads very slowly, which makes impossible a high comprehension score — although he did only two exercises incorrectly. Thus standards make it possible for a teacher to give a meaning to each score.

Interpreting scores by graphical representation. Some persons grasp the meaning of facts more easily when they are represented graphically. These standards and scores are easily represented by distances on a straight line as shown in Fig. 7. In this figure the standards for Grades 6, 7, and 8 are represented by distances on the two horizontal lines. In each case the scale has been chosen so that the sixth-grade standard for rate (92) is directly under the comprehension standard (21). The same has been done for the seventh- and eighth-grade standards. This plan produces an irregular scale, but has the advantage that the standards for any grade lie on a vertical line. In the figure lines have been drawn to represent the scores of four pupils given in Table V.

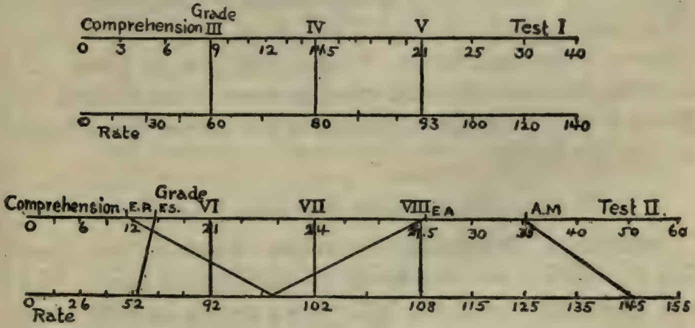


Fig. 7. Showing a Scheme for the Graphical Representation of the Scores of Monroe's Standardized Silent Reading Tests. Lower Figure shows the Scores of Four Seventh-Grade Pupils.

In the case of pupil E. A., for example, a glance at the figure tells us that he reads silently with eighth-grade ability, but his rate is slightly less than seventh-grade standard. Pupil E. R. reads as rapidly, but is conspicuously below sixth-grade standard in comprehension. One advantage of this plan of graphical representation is that it gives a meaning to the amount of difference between the score and the standard. It means more to say that a pupil is two grades below standard than to say his score is 21 when the standard is 27.5.

Interpreting class scores. After the scores of the pupils in a class have been recorded on the class record sheet, the median score of each distribution is found. The median scores are the "class scores." (See page 29.) These are interpreted in the same way as the scores of individual pupils. However, one should remember that since the median represents the "average" or general status of the group, deviations from the standards will not be so large as the deviations of individual scores. Thus a difference of a few units between a standard and a median score is much more significant than a similar difference in the case of the scores of individual pupils.

II. COURTIS'S SILENT READING TEST NO. 2

Standards. The standard scores for Courtis's Silent Reading Test No. 2 are printed on the class record sheet. They are reproduced in Table VII. They are to be used in the same general way as the standards for Monroe's Standardized Silent Reading Tests in the interpreting of both individual and class scores. These standards represent the performance of the "average" or typical pupil in the respective grades. The plan of graphical representation described above may profitably be used here also.

TABLE VII. STANDARD SCORES FOR COURTIS'S SILENT READING TEST NO. 2

Grade	II	III	IV	V	VI
Words per minute	84	113	145	168	191
Questions in five minutes	16	24	30	37	40
Index of comprehension	59	78	89	93	95

Interpreting individual scores. In Folder D, Series R, Courtis gives the following suggestions for interpreting pupils' scores obtained by using his test.

> Three types of comprehension scores are possible.
>
> (A) Large negative indices.
> (B) Zero, small positive, or negative indices.
> (C) Large positive indices.

The general meaning of these is as follows:

(A) The child misreads. That is, he not only fails to comprehend what he reads, but he persistently gets the opposite meaning from that in the sentence.

(B) The child is guessing at the answers and is not reading at all. Repeat the test with appropriate explanations until you are sure he understands what is wanted. Then measure him again, using a new form of the test. Two forms have been printed: The Kitten Who Played May Queen (Form I), and The Kitten Who Went to a Picnic (Form II). Order by form number.

(C) All other scores are to be interpreted in the light of the relation between rate of reading and the rate of answering questions. The general scheme is as follows:

High and *low* mean higher or lower than the median score of the class.

Scores				*Interpretation*
Type	*Rate of reading*	*Rate of answering questions*	*Index of comprehension*	*Probable meaning*
1	High	High	High	Marked ability.
2	High	High	Low	Needs training in accuracy.
3	High	Low	High	Defect in mechanical skill offset by intelligent re-reading until meaning is comprehended.
4	High	Low	Low	Poor training or poor ability.
5	Low	High	High	Cautious, careful reading on first trial. Such children usually make much higher scores on second trial.
6	Low	High	Low	Marked lack of intelligence.
7	Low	Low	High	Lack of native ability, but good training.
8	Low	Low	Low	Lack of native ability, or marked defects in training.

Interpreting class scores. To assist one in interpreting the class scores of a building or of a school system, Courtis has devised the graph sheet shown in Fig. 8. This device shows in a very effective way the median scores for "Questions answered" and "Index of comprehension" of an entire building. The circles through which the dotted line passes represent the standard scores. They are joined by the line so as to aid one in comparing their position with the position of the scores of any class. The directions for drawing this graph are reproduced just below the figure. The position of

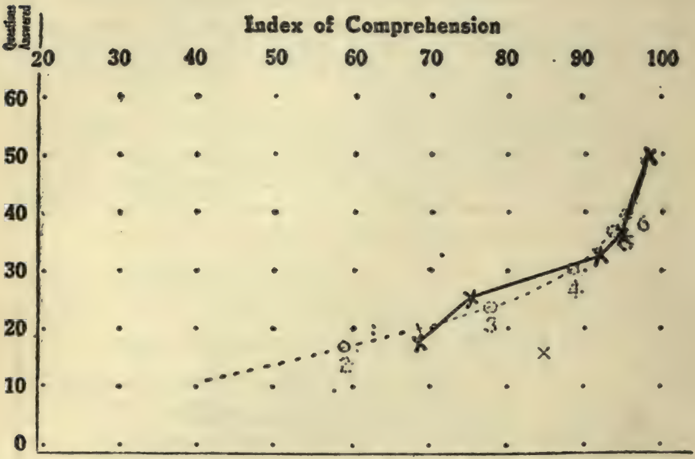


Fig. 8. Showing graphically the Median Scores of a School in Silent Reading as determined by the Courtis Silent Reading Test No. 2. (*Table VIII.*)

For each class, move a pencil up, but not touching, the scale for questions answered at the left of the figure until a point is reached corresponding to the class score for questions answered. Then move the pencil parallel to the scale for index until a point is reached which is directly below the point on the scale corresponding to the index for the class. Finally lower the pencil to the paper at this point and make an X. Join each X to the next by a straight line.

In the figure an X has been drawn to represent the scores, 15 questions answered and 85 per cent index.

the X's through which the solid line passes represents the scores given in Table VIII. Among the things which this figure tells us the most obvious are: (1) The second grade is noticeably above standard in "Index of comprehension," but slightly below in number of questions answered; (2) the sixth grade is above in both abilities, particularly in number of questions answered; (3) the third, fourth, and fifth grades are near standard; the fifth grade is exactly standard.

TABLE VIII. THE SCORES OF ONE SCHOOL

Grade	II	III	IV	V	VI
Questions answered	16	26	32	36	50
Index of comprehension	69	76	92	94	98

III. THORNDIKE'S VISUAL VOCABULARY SCALE

Standards. This scale has not been satisfactorily standardized, but we give in Table IX the average score for eighteen cities in Indiana[1] and for Louisville, Kentucky.[2] These scores are based on the use of another vocabulary scale which is supposed to be equal in difficulty to the one described on page 36. Thus the facts of Table IX may be used as tentative standards for interpreting individual and class scores.

TABLE IX. MEDIAN SCORES IN VISUAL VOCABULARY
(THORNDIKE SCALE A)

Grade	III	IV	V	VI	VII	VIII
Eighteen Indiana cities	4.00	5.26	6.00	6.66	7.29	7.91
Louisville		4.4	5.3	6.4	7.1	8.2

IV. GRAY'S ORAL READING TEST

Standards. The standards for this test are given in Table X. It will be noticed that after the third grade the increase in the standards from grade to grade is only one unit, except in the seventh, where there is a decrease from the sixth. This condition is caused by the particular way in which the scores are computed, and does not mean that a pupil reads orally

---

[1] Haggerty, M. E., *The Ability to Read: Its Measurement and Some Factors Conditioning It.* Indiana University Studies, vol. IV, no. 34. (January, 1917.)

[2] Race, Henriette V., "The Work of a Psychological Laboratory, *Educational Administration and Supervision,* September, 1917.

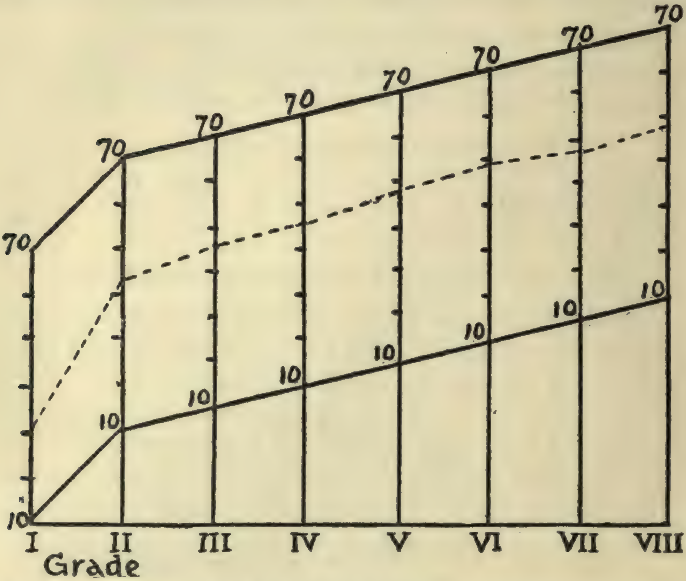


FIG. 9. SHOWING A SCHEME FOR THE GRAPHICAL REPRESENTATION OF SCORES ON GRAY'S ORAL READING TEST. (*After Gray.*)

no better in the eighth grade than he did in the fifth. This apparent inconsistency in the scores for the several grades is corrected in the plan of graphical representation shown in Fig. 9. The vertical line for each grade has a scale which begins at a different height. The position of the broken line represents the standard scores. This diagram may be used for interpreting either individual or class scores.

TABLE X. STANDARD SCORES FOR GRAY'S ORAL READING TEST

Grade	I	II	III	IV	V	VI	VII	VIII
Standard	31	43	46	47	48	49	47	48

V. CORRECTING DEFECTS IN SILENT READING

Scores furnish a basis for improving instruction. In order that the greatest benefit may be derived from the use of standardized tests, it is important that those using them understand the purpose of such tests. Their function is to furnish reliable information concerning what pupils are able to do in a certain field, such as silent reading or oral reading. The mere giving of the tests does not increase the abilities of the pupils, but when a teacher knows the abilities of his pupils and the standard scores for their grade, he has information which will be very helpful in planning future instruction. In the following pages we give records of the scores of a few typical classes and suggestions for improving the conditions represented by these scores. The first three illustrate scores obtained by using Monroe's Standardized Silent Reading Tests and the next two illustrate scores obtained by using Courtis's Silent Reading Test No. 2.

Type I. Below standard in comprehension. The first type which we shall consider is that of a class which is conspicuously below standard in comprehension as shown by Monroe's Standardized Silent Reading Tests. The scores of such a fifth-grade class are shown in Fig. 10. The fifth-grade standards are: Rate 93, Comprehension 21.0. The intervals in which these standards fall are indicated in Fig. 10. The class score for rate is slightly below standard, but this is a minor matter compared with the position of the comprehension scores.

Individual differences. Another very noticeable feature of Fig. 10 is that the scores are widely scattered, which means that the pupils of this class which have been grouped together for instruction do not possess equal ability in silent reading. In fact they are shown to differ very widely. This condition is not unusual, but the pupils of some classes are found to be more closely grouped than others, and we may say that while it probably is not possible to eliminate individual differences altogether, a closely grouped set of scores is one "earmark" of good teaching. Individual differences will be considered more fully under Type III.

Rate Score		**Comprehension Score**	
Interval	**Number of Pupils**	**Interval**	**Number of Pupils**
Above 130			
126 to 130	1		
121 to 125			
116 to 120			
111 to 115		80 & above	
106 to 110	2	70 to 79.9	
101 to 105		60 to 69.9	
96 to 100	2	50 to 59.9	
91 to 95		45 to 49.9	
86 to 90	4	40 to 44.9	
81 to 85	2	35 to 39.9	
76 to 80		30 to 34.9	
71 to 75		27 to 29.9	
66 to 70	1	24 to 26.9	
61 to 65		21 to 23.9	1
56 to 60	2	18 to 20.9	1
51 to 55		15 to 17.9	3
46 to 50		13 to 14.9	1
41 to 45		11 to 12.9	3
36 to 40		9 to 10.9	2
31 to 35		7 to 8.9	1
26 to 30		5 to 6.9	
21 to 25		4 to 4.9	1
16 to 20		3 to 3.9	
11 to 15		2 to 2.9	1
6 to 10		1 to 1.9	
0 to 5		0 to .9	
Total	14	Total	14
Median	89	Median	12.7

FIG. 10. SHOWING THE SCORES OF A FIFTH-GRADE CLASS ON MONROE'S STANDARDIZED SILENT READING TESTS. (*Type I.*)

The causes of a lack of comprehension. The teacher of the class whose scores are given in Fig. 10 faces the problem of improving their ability to comprehend. The lack of comprehension may be due to one or more of several causes: (1) A lack of a good "method" of reading silently. (2) A lack of practice in reading silently with care due (*a*) to insufficient opportunity or (*b*) to the absence of a strong motive. (3) Not sufficiently acquainted with the vocabulary. (4) Miscellaneous causes, such as becoming confused on the test or failing to understand what is to be done. If the teacher exercises care to follow the directions in giving the test such causes as this are unlikely to happen for the entire class and hence do not need to be considered in this place.

Diagnosis, or locating the cause of poor comprehension. (*a*) *Vocabulary.* In the case of a given class it will be necessary for the teacher to determine which of the above causes apply. He will frequently be able to do this simply by reason of his acquaintance with the pupils. If he is doubtful, Thorndike's Visual Vocabulary Test may be used to determine if the poor comprehension is due to the lack of vocabulary, or, if this test is not available, the teacher may select a list of words from the exercises read and ask the pupils to define them and to use them in sentences. This may be done either orally or in writing.

(*b*) *Method.* Evidence of a poor "method" of reading will be found in the character of the pupil's responses to the exercises. Efficient silent reading involves three steps: (1) assigning to each word or phrase its correct meaning; (2) combining the several elements of meaning, giving to each its proper weight or significance; (3) verifying or comparing the meaning (in the case of Monroe's Standardized Silent Reading Tests, the answer to the question) with the sentence or exercise to see if it is the correct meaning. Some pupils do not go through these steps. They merely "fish around" in the exercise for a word or phrase to use as the answer to the question. This is not reading; it is only guessing and not a "method" of reading. Evidence of this procedure will be found in the pupil's responses. If they are uniformly unreasonable or absurd, it is reasonably certain that the pupil is "guessing" or writing down the first thing which comes into his mind unless he is very deficient in vocabulary.

An illustration of "guessing" in silent reading. Sometimes it is wise to secure further evidence. This can be done by requiring the pupil to answer questions based upon a paragraph such as the following:[1]

In Franklin, attendance upon school is required of every child between the ages of seven and fourteen on every day when school is in session unless the child is so ill as to be unable to go to school, or some person in his house is ill with a contagious disease, or the roads are impassable.

1. What is the general topic of the paragraph? ....
2. How many causes are stated which make absence excusable? ....
3. What kind of illness may permit a boy to stay away from school, even though he is not sick himself? ....
4. What condition in a pupil would justify his non-attendance? ....

The following answers to the above questions by sixth-grade pupils are taken from a report by Thorndike[2] and are typical of responses which indicate that the pupil is "guessing" at the answer to the question, either on the basis of what the paragraph or some word in it suggests to him or on the basis of his general experience. The number following the answer is the number of times it occurred per hundred papers. (Two hundred papers were examined.)

---

[1] This paragraph and the questions are taken from Thorndike's Scale Alpha 2, Part II, Set V. This scale may be purchased from the Bureau of Publications, Teachers College, New York City. A similar test, arranged in a more convenient form for classroom use and called the Minnesota Scale Beta, may be secured from the Bureau of Coöperative Research, University of Minnesota, Minneapolis, Minnesota.

[2] *Journal of Educational Psychology*, vol. 8, p. 324, June, 1917.

Question 1.

Franklin	4 1/2
Franklin attends to his school	1/2
It was a great inventor	1/2
Because it's a great invention	1/2

Question 2.

If the child is ill	2
Illness.	1
Very ill.	3
An excuse	2

Question 3.

If Mother is ill	5 1/2
Headache, ill	1/2
A sore neck	1/2
When a baby is sick	1/2
When the roads cannot be used	1/2

Question 4

By bringing a note	6
To have a certificate from a doctor that the disease is all over	1/2
Torn shoes	1/2
When he acts as if he is innocent	1/2
Being good	1/2
Get up early	1/2
Come to school	1 1/2
If he lost his lessons	1/2
Truant	1
If some one at his house has a contagious disease	6 1/2
Not smart	1/2
By not staying home or playing hookey	1/2

An illustration of failure to verify meaning. In some cases the answers to questions indicate that the pupils are not "guessing," but are inaccurate because they fail to verify or compare their answer with the paragraph read. Obviously this step was not taken by the pupils who made the answers quoted above, but they committed another error. They did not try to read. They simply "guessed" or took the first idea which came into their minds and did not even ask if it was sensible or foolish. But such answers as the following suggest that the pupil "tried" to read but failed to answer correctly, partly because he did not verify his answer:

Question 1.	
The attendance of the children	1/2
School	7 1/2
About school	4
How old a child should be	1/2
Question 3.	
Serious	1/2
Contagious disease, roads impassable	1 1/2
Question 4.	
Somebody else must have a bad disease	1/2
Illness, lateness, or truancy	1/2

Thorndike says:

> Reading may be wrong or inadequate because of failure to treat the responses made as provisional and to inspect, welcome, and reject them as they appear. Many of the very pupils who gave wrong responses to the questions would respond correctly if confronted with them in the following form:
>
> Is this foolish or is it not?
> The day when a girl should not go to school is the day when school is in session.
> The day when a girl should not go to school is the beginning of the term.
> The day etc......... is Monday.
> The day is fourteen years.
> The day is age eleven.
> The day is a very bad throat.
> Impassable roads are a kind of illness.
> He cannot pass the ball is a kind of illness.
>
> They do not, however, of their own accord test their responses by thinking out their subtler or more remote implications. Even

very gross violations against common sense are occasionally passed.[1]

Reason for failure to verify meaning. In another place he comments upon the general reason for this:

> There seems to be a strong tendency in human nature to accept as satisfactory whatever ideas arise quickly — to trust any course of thought that runs along fluently. If the question makes the pupil think of anything or if he finds anything in the paragraph that seems to belong with the question, he accepts it without criticism. Wrong answers are, in reading tests with all ages, too frequent in comparison with admissions of ignorance. This holds of tests in other subjects also.
>
> It seems probable that in scoring pupils' work in schools an admission of ignorance should not be penalized as heavily as an absurd or specially harmful error, and that inadequacies and errors in general should be penalized somewhat more heavily than they now are, at least in the many cases where it is much more useful to know that one does not know and to say so, than to respond wrongly. On the other hand, a mere chronic suspicion and skepticism concerning one's ideas is undesirable. It is healthy to trust the ideas which the laws of habit produce, provided we maintain an active watch for other ideas which may tell whether the first ones are appropriate. The pupil should learn to criticize his responses, but not to be frightened into a mental paralysis.[2]

An illustration of the lack of vocabulary. The lack of vocabulary is indicated by such responses to the first question of the exercise on page 54 as the following:

Question 1.	*Per cents*
A few sentences	1/2
Made of complete sentences	1/2
A sentence that made sense	1/2
Subject and predicate	1/2
A letter	1/2
Capital	5 1/2

---

[1] *Journal of Educational Psychology* (June, 1917), p. 330.

[2] *Elementary School Journal* (October, 1917), vol. 18, p. 107.

	Per cents
A capital letter	1/2
The first word	1/2
Leave half an inch space	2 1/2
The heading	1/2
Period	1/2
An inch and a half	1/2
An inch and a half capital letter	1/2

How to correct such defects. 1. Motivation. Pupils who are "reading" in the ways described on the preceding pages must, *first*, be caused to desire to read better, that is, their silent reading must be motivated more strongly; *second*, they must be given practice in careful reading by the teacher making use of and creating situations in which emphasis is upon thought-getting and not upon oral expression or rate of reading.

Silent reading motivated by the use of standardized tests. It has been the experience of many teachers who have used standardized tests that a strong motive is frequently created by telling the pupils the standards for their grade and the scores of their class. This gives the pupils a definite aim to work for and a statement of the progress which the class must make. It secures for the teacher the coöperation of the class, which is very important. The writer has visited classrooms where the teacher had the class scores and the standards represented graphically on a chart which was posted in the front of the room. If the class was below standard, the pupils were interested in having the class scores brought up to standard.

Commendable results have also been secured by having each pupil compare his scores with the standards. This stimulates the pupil to compete with an objective standard and not with his classmates. Thus the undesirable feature of competition is eliminated. If the tests are repeated from time to time, the pupil also has the advantage of comparing his successive scores. He thus learns the amount of his progress. The teacher should bear in mind that probably all pupils will not attain the standards and that some will exceed them. A pupil who is below standard, but is making progress, may be doing all that is possible for him in the time that is devoted to reading. If he is, the teacher must make certain that he does not become discouraged.

2. Emphasis upon thought-getting. (*a*) *In the primary grades.* Children in the primary grades should from the start have exercises in which the meaning is the only significant element, and the response is not in terms of words said, but things done, or interpretations made. For example, let it be the usual thing for the child to carry out the directions contained in the word or sentence. The primary teacher should be supplied with some hundreds of cards upon which such sentences or short paragraphs as the following are printed or written:

(1) Draw a picture of a flag on the blackboard.
(2) Make a sound like a cross kitty makes when a dog chases her.
(3) Hide behind the door.
(4) Play that you are carrying a cup full of water and do not wish to spill any of it.

These cards should be graded in such a way that certain ones will contain only the words taught in the first reading lessons. As more words are learned, more cards will become available. Variety in handling the exercises may be introduced in scores of ways which will readily occur to a resourceful primary teacher. Many other devices having the same aim will also occur to the teacher. The essential thing is that practice in translating written or printed language into action instead of words should be started early, thus producing the habit of advancing through a paragraph by thought-units rather than by letters, syllables, or words.

(*b*) *Above the primary grades.* In grades above the primary the problem is fundamentally the same as stated for the primary, but the devices must vary.

First, whenever reading is done orally, be sure that what the child is reading is new to most of his listeners. Be sure, too, that the other pupils are listening, and not following along with the reader in another copy of the same book. No method of reading is more faulty in intermediate grades than that in which other members of the class are watching for a word error of the reader, ready to call attention at once to such a mechanical mistake. This method centers the attention of the reader constantly upon the mechanics and never develops the habit of attending first to the thought. Whereas, if the reader realizes that his hearers know nothing of the content of his selection except what they gather from his reading, then giving the thought instead of pronouncing the words becomes the controlling factor in his consciousness. It follows from this that only selections, the thoughts in which are vital to children, should be used as subject-matter for such reading. Then let the one who has read such a selection defend the selection against questions or criticisms of the class. In short, center attention upon the meaning, even at the expense, if necessary, of accuracy in pronunciation, enunciation, and expression.

Second, let the amount of reading which is compellingly interesting be increased. Supplementary reading in geography, history, science, and literature should be given a larger place. Require that the reports made upon such readings be rather exact, but let the selections be reasonably easy for the children. Gain in facility in silent reading cannot be secured by holding the children to selections which are so difficult that word-troubles absorb all the attention. One must be able to go with ease through the successive thoughts before the habit of attending to the thought can be acquired.

Third, make all the industrial and playground exercises give a far greater measure of service in teaching reading than they now commonly give. How singularly short-sighted we are to ask a child to follow the directions printed in his arithmetic for finding the per cent that one number is of another, but employ a teacher to give orally the directions for playing a new game, making a raffia basket, or planting beans. The very things which come nearest the natural interests of the children, concerning which they would most zealously read if they had the paragraphs containing the needed directions, are given to them orally. When interesting school exercises require a careful following of directions, then those directions make the most effective silent reading material. But in practice we seldom make use of them. This fault is due to a failure to understand the distinction between the aim of the intermediate grades and the aim of the upper grades. If we realized that all the work of the intermediate grades should be made to develop skill in using the tools of learning, then we should not conduct these exercises without making them aid in teaching reading.

(*c*) *In the upper grades.* Passing now to the situation presented when the score of a class above intermediate grades is found to be low, we have the most serious task of all. The junior high-school or upper-grade pupil should be able to proceed with his school tasks without much attention to the tools he is using. It is not the primary function of this department of the school system to increase the children's facility in the handling of these tools. However, success in nearly all the tasks undertaken in the upper grades depends upon the skill which the children are expected to possess in the tool subjects. A compromise is, therefore, necessary, if children in the junior high school or seventh and eighth grades, are found deficient in their ability to read silently. A few suggestions are here offered in the hope that some help may come from them, although it is realized that correcting reading faults at this stage is very difficult.

First of all, the children's own conscious efforts should be obtained in the direction of correcting the faults. Then, too, the teacher should see that he is observing the same fundamental principles stated for the intermediate grades. Comprehension, and not mechanics, must be made the test of all reading, whether in history, science, or literature. The material selected for use must be sufficiently easy so that the children are not tied up in word or language difficulties. Again, to overcome the habit of proceeding by too small units, practice must be afforded in advancing by short sentences or phrases.

In case the trouble seems to be that the children read fluently enough orally, but get little of the thought, introduce a great deal of the sort of reading requiring close attention to the thought. For example, use rule books for football, basket-ball, and the like for those interested in games; catalogue descriptions; directions for making certain stitches; the more involved arithmetic problems; and so on. These things possess a minimum of word-difficulty and a maximum of thought-difficulty. They require the imagination to construct a picture little by little and hold it up for constant modification as the reading proceeds. Thus, attention is focused on thought.

Where the class appears to have the right habits of reading silently, but have had insufficient practice, the obvious suggestion is to give them all the practice possible. Much supplementary reading upon which they make only meager reports, if any, will help. Try to secure as much general home reading as possible. See that an abundance of interesting things is available for reading, and stimulate interest by having the children's criticisms of them given before the class.

3. **Exercises requiring careful reading to answer ques-**

tions. Exercises of the kind shown on page 54 can be used to an advantage in teaching pupils to comprehend what they read. Exercises may be taken from the tests mentioned in the footnote, but a teacher will not find it difficult to construct similar exercises by asking a series of questions based upon paragraphs in the pupils' geography, history, or other texts. If supplementary readers are available, they can also be used in this way. The teacher can write the questions on the board or dictate them. The next day the papers should be returned and the attention of the pupils called to their errors. This plan can be varied by having the pupils turn to a particular paragraph in their text and prepare an appropriate set of questions on it. The teacher can judge the questions upon the basis of whether they call for the important ideas in the paragraph.

The use of such exercises does two things: first, answering the questions will give the pupils an idea of what careful reading involves; second, their attention will be directed to the necessity of verifying their answers.

4. **More attention to vocabulary.** If the cause is found to be a lack of acquaintance with the meaning of the words used, more attention should be given to vocabulary. In the upper grades the use of the dictionary will help, but the most important thing is that the teacher shall definitely recognize the necessity for teaching the meaning of words, not merely formal dictionary definitions, but rich, comprehensive meanings which are directly connected with the experiences of pupils. It is frequently worth while to spend five or ten minutes in a class discussion of the meaning of an important word. The use of a vocabulary test will tend to direct the attention of the teacher to the necessity for doing this. It may also happen that when the pupil finds that he is below standard in vocabulary his coöperation will be secured.

Rate Score		Comprehension Score	
Interval	Number of Pupils	Interval	Number of Pupils
Above 130			
126 to 130			
121 to 125			
116 to 120			
111 to 115		80 & above	
106 to 110	1	70 to 79.9	
101 to 105		60 to 69.9	
96 to 100		50 to 59.9	
91 to 95		45 to 49.9	
86 to 90		40 to 44.9	
81 to 85	3	35 to 39.9	
76 to 80		30 to 34.9	
71 to 75		27 to 29.9	
66 to 70	3	24 to 26.9	
61 to 65		21 to 23.9	1
56 to 60	4	18 to 20.9	
51 to 55	5	15 to 17.9	
46 to 50		13 to 14.9	
41 to 45	9	11 to 12.9	2
36 to 40	3	9 to 10.9	2
31 to 35	3	7 to 8.9	10
26 to 30		5 to 6.9	3
21 to 25	1	4 to 4.9	4
16 to 20	1	3 to 3.9	2
11 to 15		2 to 2.9	
6 to 10		1 to 1.9	5
0 to 5	1	0 to .9	5
Total	34	Total	34
Median	43	Median	5.9

FIG. 11. SHOWING THE SCORES OF A FOURTH-GRADE CLASS ON MONROE'S STANDARDIZED SILENT READING TESTS. (*Type II.*)

5. Providing opportunity for practice in silent reading. This point will be discussed more fully under Type III, but attention should be called to this means of improving the ability of pupils to comprehend what they read. Reading is an art and pupils must have much practice. Supplementary reading material of the right kinds should be provided and definite provision should be made for opportunity to read it during school hours. The teacher should look upon supplementary reading as an important school activity and one requiring his supervision.

Summary for Type I. Under this type we have considered the case of a class which has a low comprehension score. The causes considered for this condition are: (1) failure to use a good "method" of reading; (2) a lack of practice; and (3) insufficient vocabulary. We have suggested plans for diagnosis or locating the cause and have given several typical illustrations of the causes mentioned. The methods of correcting these defects have been presented under these heads: (1) motivation; (2) emphasis upon thought-getting; (3) exercises requiring careful reading to answer questions; (4) attention to vocabulary; (5) providing practice in silent reading. These methods will be considered again under Type III as means of correcting individual defects.

Type II. Below standard in rate of reading. In Fig. 11 there is shown the record of a fourth-grade class which reads very slowly. The pupils also made low scores on comprehension but this is due in part to their slow rate of reading because when Monroe's Standardized Silent Reading Tests are used a high comprehension score is impossible for slow readers.

Causes of slow reading. Three causes may be given for a situation such as is illustrated in this second type: (1) the common belief that in order to read well one must read slowly; (2) over-emphasis upon oral reading which results in the pupil pronouncing the words to himself when he reads silently; (3) failure on the part of the teacher to recognize that the rate of reading is important.

How to increase the rate of silent reading. 1. Motivation. One effective plan is to furnish a strong motive. This can be done by using standardized tests as suggested on page 58. Interesting stories or references for supplementary reading will often be effective. If a pupil becomes interested in a story, either by having had a part of it read to him or by having read the first of it himself, he will be anxious to read the rest of it to "see how it comes out." While the quality of the reading should not be neglected, the emphasis should be on the rate of reading. In order that this may be done, the reading material must be simple.

2. Emphasizing rate of silent reading by informal testing. One reason why pupils read slowly is that the teacher pays no attention to the rate of silent reading. In Chapter I we pointed out that one defect in our ordinary measurement of results was the neglect of the rate of work. The rate of reading is an important "dimension" of the ability to read silently. In many cases a teacher can increase the rate of his pupils' reading by simply recognizing it as one "dimension" of the ability to read. This can be done by asking the pupils to read silently beginning with a certain paragraph in their text (school reader, geography, history, or elementary science). At the end of a suitable period, three to five minutes, stop them and have them count the number of lines read. This number will be a crude measure of the rate of reading. This should be a part of the regular instruction in silent reading. If the teacher doubts the quality of the reading, it can be tested informally by having the pupils answer a set of questions based upon the lines read.

In the survey of the Cleveland Public Schools an informal silent reading test was given by having the pupils read silently in the Jones Readers. After some preliminary testing to give the pupils an understanding of what they were to do, the teacher read aloud a page to the class, the pupils having their books open. When he came to the turning of the page the teacher stopped reading and noted the time. The pupils continued the reading silently. At the end of *one* minute they were stopped and the number of lines read were counted. The pages used for the test and the average number of lines read are given in Table XI. In interpreting the average number of lines given in this table, one must remember that the material read in the upper grades was more difficult and that the lines contained more words. He may use these facts as tentative standards for judging his pupils when testing their rate of reading in the way suggested.

TABLE XI.* SHOWING RATE OF SILENT READING IN INFORMAL TESTING

Grade	*Book*	*Preliminary page*	*Test pages*	*Average number of lines read — 44 schools*
2A	II	101	102–103	16
3A	III	97	98– 99	22
4A	IV	61	62– 63	21
5A	V	47	48– 49	20
6A	VI	63	64– 66	24
7A	VII	63	64– 66	21
8A	VIII	247	248–249	21

* Judd, C. H., "Measuring the Work of the Public Schools," *Cleveland Education Survey*, p. 261.

Rapid readers good readers. It has commonly been thought that a thorough understanding required that the pupil should read slowly and carefully, and that the rapid reader understood very little of what he read. It is, of course, true that the pupil who reads with extreme rapidity, or "skims" over the page, does not comprehend completely what he reads, but we now have evidence which shows that in many cases a rapid reader is a "good" reader and a slow reader is a "poor" reader.

Fig. 12 is reproduced from the *Report of the Cleveland Survey* to show the relation which was found to exist between rate and quality of silent reading as measured by Gray's Silent Reading Tests.[1] On the basis of their scores 1831 pupils were divided into the nine groups indicated in the figure.

---

[1] These tests are not described in this book because they are not suited to general classroom use. For a complete description of them the reader is referred to Gray, William S., *Studies in Elementary School Reading Through Standardized Tests.* (Supplementary Educational Monographs, University of Chicago Press.)

The per cent of the pupils in each group is given by the number inside of the circle. The size of the circle represents the size of the group. The figure shows very clearly that

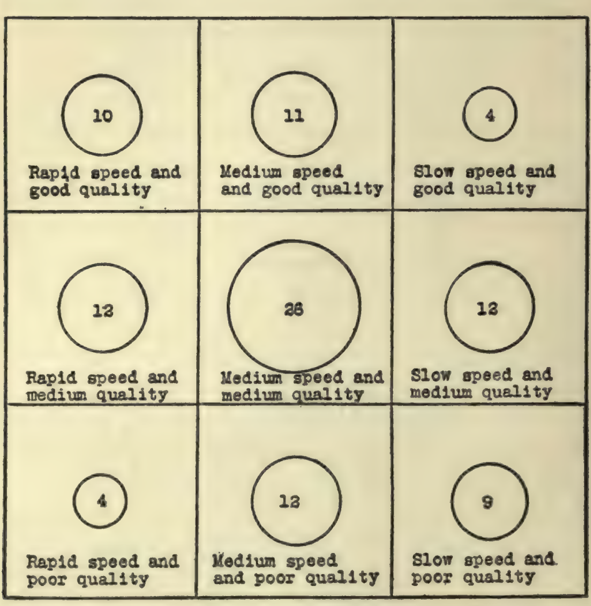

10
Rapid speed and good quality
11
Medium speed and good quality
4
Slow speed and good quality
12
Rapid speed and medium quality
28
Medium speed and medium quality
12
Slow speed and medium quality
4
Rapid speed and poor quality
12
Medium speed and poor quality
9
Slow speed and poor quality

Fig. 12. Per cent of 1831 Cleveland Pupils found in each on Nine Speed and Quality Groups in Silent Reading. (*From Judd's "Measuring the Work of the Public Schools."*)

a rapid reader is good in quality more frequently than he is poor in quality and the opposite is true for slow readers. The application of this fact is that in many cases pupils will improve in quality of reading when they increase their rate. Teachers should not expect to secure a higher degree of comprehension by urging their pupils to read more slowly.

3. **More opportunity for silent reading.** Practice in reading is even more necessary for producing the ability to read

rapidly than for engendering the ability to comprehend. The suggestions on page 64 apply here also. It is particularly important that the material should not be difficult to understand.

4. **Less emphasis upon oral reading in the intermediate and grammar grades.** Oral reading is necessary in the primary grades, but as the pupil progresses from grade to grade, more emphasis should be placed upon silent reading and less upon oral reading. From about the fourth grade silent reading should receive the greater emphasis. Failure to do this frequently causes the pupil to acquire habits which make rapid silent reading impossible. Two cases of slow readers due to this cause are described on pages 73 and 84.

Summary for Type II. Under this type we have considered the case of a class which reads too slowly. The following causes were considered: (1) attempt to secure a high degree of comprehension by urging the pupils to read more slowly; (2) over-emphasis upon oral reading; (3) failure of teacher to recognize the rate as important. For increasing the rate of silent reading the following correctives were given: (1) motivation; (2) greater emphasis upon the rate of reading; (3) more opportunity for silent reading; and (4) less emphasis upon oral reading.

Type III. Scores too widely distributed. In Fig. 13 we show the scores of a fifth-grade class whose median scores are approximately standard. (The test was given February 4, and hence it is not to be expected that the class had attained the seventh-grade May standard.) The noticeable thing about this record is that the scores for rate range from one score between 31 and 35 to one above 130, and for comprehension from one score between 1 and 2.9 to one between 35 to 39.9. Thus, there are in this fifth-grade class some pupils below third-grade standards (60, 9.0) and others above the eighth-grade[1] standards (108, 27.5). The teacher's particular problem in this case is with the pupils who have the low scores. Those who are above standard do not constitute a problem, if they are above in both rate and comprehension, except that the teacher should consider whether these pupils could spend the time now devoted to reading more profitably on some other subject. It might happen, for example, that some of these pupils might be below standard in arithmetic, spelling, or

Rate Score		**Comprehension Score**	
Interval	**Number of Pupils**	**Interval**	**Number of Pupils**
Above 130	1		
126 to 130	1		
121 to 125			
116 to 120	2		
111 to 115		80 & above	
106 to 110	4	70 to 79.9	
101 to 105		60 to 69.9	
96 to 100	1	50 to 59.9	
91 to 95		45 to 49.9	
86 to 90	2	40 to 44.9	
81 to 85		35 to 39.9	1
76 to 80		30 to 34.9	2
71 to 75		27 to 29.9	2
66 to 70	1	24 to 26.9	1
61 to 65		21 to 23.9	4
56 to 60	2	18 to 20.9	
51 to 55	1	15 to 17.9	
46 to 50		13 to 14.9	
41 to 45	2	11 to 12.9	1
36 to 40		9 to 10.9	1
31 to 35	1	7 to 8.9	1
26 to 30		5 to 6.9	4
21 to 25		4 to 4.9	
16 to 20		3 to 3.9	
11 to 15		2 to 2.9	1
6 to 10		1 to 1.9	
0 to 5		0 to .9	
Total	18	Total	18
Median	93	Median	21.8

FIG. 13. SHOWING THE SCORES OF A FIFTH-GRADE CLASS ON MONROE'S STANDARDIZED SILENT READING TESTS. (*Type III.*)

---

[1] Accurate comparison of fifth-grade scores with eighth-grade standards is not possible because different tests are used in these grades. However, the statement is probably true.

language. If so they need to devote some extra time to these subjects.

The condition shown in Fig. 13 may be due to an unwise classification of the pupils and some adjustments should be made which would reduce the wide range of scores. However, this would probably reduce the number of cases only slightly and the problem would still remain.

Uniformity in instruction for all the members of a class widens variability among them, making the weak ones relatively weaker and the strong ones relatively stronger. To prevent this widening of the variability more attention must be given to individual instruction. This does not mean a leveling of all members of a class, but rather affording a maximum of opportunity to each member to do those things most needful to him. Those things which he can already do well he should not be required to do, even though some other members of the class need to do them.

Those children falling far below the median of the class should be given special physical examination to discover if possible the cause. Sometimes eyesight is found to be poor. Frequently some other physical defect has prevented normal mental growth. Sometimes an examination by means of approved intelligence tests, such as the Binet-Simon tests,[1] reveals that the child is mentally incapable of doing work of the regular school type.

Illustrations of individual defects. It may be that the pupil needs to be *taught how to read silently*. Not very much attention is given to teaching pupils how to read silently. The instruction in reading is confined largely to oral reading. A pupil who has not learned how to read silently needs instruction. One teacher writes of a certain pupil as follows:

---

[1] See especially Terman, L. M., *The Measurement of Intelligence* (Houghton Mifflin Company, Boston, 1916). A simple guide for the use of the intelligence scale.

> From the tests given and from her work in English which I have had for two years, I find that she has only a vague hazy kind of meaning for many of the words needed for seventh-grade work. She does not see words in their relation to others in the sentence. When she finds a name for a combination of letters she is satisfied, thinking that she is reading. She has failed this year. I hope this may rouse her to the effort of which I am sure she is capable. If I can only make her see that reading means more than naming words and persuade her to work, I am sure she can overcome her difficulties.

Another difficulty may be vocabulary. A boy who made a low score on the silent reading test was given the vocabulary test, on which he made a very low score. His teacher describes him as follows:

> His greatest difficulty seems to be a lack of vocabulary. He memorizes history instead of studying for the thought. Lately, he has gotten away from this to some extent and begins to sum up the thought rather than repeat words when called upon. He still (after some individual instruction) finds so many unfamiliar words in any new paragraph that his progress is very slow, but he attacks his problem with more intelligence than he showed at first.

How to bring individual pupils up to standard. If the child is nearly normal physically and mentally, but has not developed ability to get meaning from printed language, he presents a problem in instruction calling for the best professional skill to solve. In dealing with such pupils the suggestion given for Types I and II can be used. Giving the pupil a strong motive frequently will solve the problem.

It is quite certain that a pupil far below the median in this basic ability has never made use of printed language to secure help in satisfying his own childish desires. If possible, situations must be brought about in which his desires or plans depend for their fulfillment upon his reading. It may be, for example, that his mother or father has been in the habit of reading stories to him. If so, and he can be made to be keenly interested in a story by having a part of it read to him, he should have to read the rest himself to satisfy his desire to know the rest of the story. Possibly he would like to be the leader in an occasional nature-study excursion, but, of course, it will be expected that he look up information concerning the things they see on the trip and be able to report later to the group. That is the business of the leader. Or he might umpire the baseball game if he made sure of the rules; or assign the parts in the coming school entertainment, if he read the various parts carefully so as to be able to make a wise assignment; or score the class compositions on the basis of which was most interesting. Such a list of possible opportunities for calling into service a child's silent reading ability might be largely extended. The two things to guard against are (1) making reading a punishment and (2) confusing child need with school need. The thing to be accomplished is to give the child a chance to do something which he really wishes to do, but cannot do without reading.

The case of a slow and inefficient reader. Judd[1] gives the case of a girl in the fifth grade who was average or above in all her school subjects except reading. In this one subject she had been rated as a poor student from the first through the fourth grade. Her health was good and she had been regular in her attendance at school. With respect to reading she is described as follows:

> Reading seems to be her greatest weakness. Her fourth-grade teacher reported her as "a slow reader who reads hesitatingly and haltingly, repeating words and phrases. Her breathing is very shallow, often causing her to pause for breath in the middle of a word or phrase. Her voice is thick, heavy, and unpleasantly nasal.

---

[1] Judd, C. H., *Reading: Its Nature and Development.* Supplementary Educational Monographs (University of Chicago Press), vol. 2, no. 4, p. 82. (The paragraph headings in this and the other quotations from this monograph have been inserted by the author.)

Silent reading is particularly distasteful to her. She always settles down to it reluctantly and tardily."

From the home comes much the same story: "She has never read a story to herself, though she has several attractively illustrated children's books. She frequently, however, after eagerly studying the illustrations in a new book, begs to have the story read to her, saying, 'You read it, mother. I can't understand it very well when I read it myself.'"

This pupil was carefully tested in both oral and silent reading. In oral reading her score was 33, while the standard is 48. She made many errors, particularly mispronunciations. In silent reading, she read even more slowly than she did orally.

Observations made during the silent-reading tests showed that there was much vocalization. The reading was done in a low whisper, and difficult words, as stated above, were spelled out letter by letter. She followed the line with her finger. In one of the early practice periods, when urged to read more rapidly, she remonstrated, saying that she could not hear the words so well if she did.[1]

The correctives which were used. From the foregoing data it is evident that her difficulties in reading were due to a lack of familiarity with printed words and a lack of method of working out new or unknown word-forms. In an effort to help her overcome this handicap she was given various types of training during eighteen weeks. The first six weeks were devoted to a great deal of oral reading. The second six weeks were spent on drills in phonics and in word analysis. During the last six weeks she was given a great deal of silent reading. While each period of six weeks thus stressed some one phase of reading, all three types of work were carried along throughout the eighteen weeks. For example, oral reading was continued with less emphasis during the last twelve weeks.

The selections for oral reading were made along the line of the pupil's school interests in history and geography. These included Baldwin's *Fifty Famous Stories* and *Thirty More Famous Stories*, Harding's *Story of Europe*, Allen's *Industrial Europe*, Carpenter's *Eu-*

---

[1] This is the case of a pupil whose defect was probably caused by overemphasis upon oral reading. The pupil had never been taught to read silently. (*Author.*)

rope, "Our European Cousins Series," the Merrill and the Horace Mann *Third* and *Fourth Readers*, Tappan's *Old World Heroes*, Terry's *The New Liberty*, and Brown's *English History Stories*.

Phonics and word analysis were emphasized during the second six weeks. Various systems of phonics with some modifications to suit the particular needs were used. Words mispronounced in oral-reading lessons were worked out phonetically, and lists of words similarly pronounced were built up and reviewed from time to time. There seemed to be a gradual growth in ability to attack an unfamiliar word. In the earlier period the pupil frequently looked at the word helplessly or pronounced a known syllable, but was unable to attack it at all phonetically. She usually asked the instructor to pronounce it. Later she began immediately to sound the new word phonetically, and though sometimes making a mistake in the length of the vowel or in the position of the accent, her manner of attack indicated that she had confidence in her own ability to work it out.

Silent reading was emphasized during the last six weeks after some training in silent reading had been given throughout the first twelve weeks. For special training paragraphs or selections dealing with topics of particular interest to the pupil were used. In many instances the original selections were edited, and the words which had been used in the phonic exercises were woven into the text. Frequently before the silent reading began a question was raised, the answer to which was to be found in the text. Oral or written reproduction or a discussion of the thought of the selection usually followed the reading. It is interesting to note, in passing, that though no effort was made to reduce the vocalization so perceptible at first, it entirely disappeared except when an unusually difficult passage was encountered.

The result. One of the significant results is that mentioned in the closing sentence of the above paragraph. The pupil had learned how to read silently without pronouncing the words in a whisper. After the correctives described above had been used, the pupil was tested again in both oral and silent reading. She showed a decided gain in rate of oral reading and a reduction in the number of errors. In silent reading her rate had increased so that she now read more rapidly silently than orally, whereas before this special training the opposite was true. At the same time she made large gains in comprehension.

> Her teachers report that Case G [the pupil described above] reads with much greater ease and fluency of expression. The quality of her voice has improved and the nasal tones have almost disappeared. She seems to enjoy reading silently much more than before training. Frequently she expresses a preference for reading a passage silently, saying, "I can do it faster."

The case of a pupil deficient in vocabulary. A seventh-grade boy is described by Judd,[1] whose most striking defect in reading was a lack of word meaning. He is described as follows:

> In general school standing he is rated as a poor student, although he is given a grade of good (B) in the manual arts, music, and physical training. In all other subjects he is poor. During the past two and a half years he has received no grade higher than C in history, geography, science, literature, composition, and grammar. In this connection it is interesting to note that progress in these subjects after the fourth grade is dependent to a large degree on ability to get thought from the printed page.
>
> His teachers report him as a shy, timid boy, easily embarrassed, lacking in self-confidence and initiative in the classroom, though very energetic and responsive on the athletic field. He rarely takes part voluntarily in class discussions, and when called on to do so responds in a few brief, fragmentary sentences, badly expressed, but usually containing a thought or an idea on the topic being considered. His English teacher finds great difficulty in getting him to read with any degree of expression, for he makes no attempt to group words into thought units. He reads in a dull, monotonous tone, slurring words and phrases. When asked to tell what he has read, he reproduces a few ideas in short, scrappy sentences, for apparently he makes few associations as he reads. His teachers in history and geography explain his poor standing in their subjects as attributable to an inability to get ideas from the text. He apparently reads as rapidly silently as any in the class, but gets and retains less of the thought.

---

[1] Judd, C. H., *Reading: Its Nature and Development.* Supplementary Educational Monographs (University of Chicago Press), vol. 2, no 4, p. 106

The tests in oral and silent reading sustained the opinions given by his teachers. In the oral test he read fairly rapidly, pronouncing the words mechanically and enunciating poorly. . . .

The test in silent reading defined more clearly his apparent difficulties. . . . Clearly this particular seventh-grade boy ranks in comprehension at a lower level than the poorest readers in the two preceding grades. This result verifies the estimates of his teachers of history and geography.

A résumé of the facts brought out by the tests would seem to indicate that he has acquired a mastery of the rudimentary mechanics of word recognition, but lagged far behind in the mastery of word meaning. He read words as mere names and not as symbols of ideas.

The correctives which were used. How to build up a background of meaning that would form a basis for his reading was and still is an urgent and difficult problem. Because of his interest in animal stories and tales of camp and pioneer life, emphasis was laid throughout the eighteen weeks on literature dealing with these topics. *The Boy Scouts' Manual*, Custer's *Boots and Saddles*, Roosevelt's *Winning of the West*, Southworth's *Builders of Our Country*, Book II, the Merrill and the Horace Mann *Fourth* and *Fifth Readers*, Muir's *Stickeen*, Coffin's *Boys of '76*, the Seton Thompson and Kipling stories, and similar literature were drawn upon freely. Silent reading was continued throughout the eighteen weeks, but was especially emphasized during the first six weeks and again during the last six weeks. After reading a selection the pupil reproduced it orally or in writing. These reproductions at first were so meager and inadequate that he frequently had to re-read several items before he could answer the questions raised. Many selections were read in this way paragraph by paragraph, and the main points jotted down to assist in the organization of the thought.

Before the work had progressed very far it became apparent that definite word-study was necessary in order to build up a background of meaning. Words were studied in the context for meaning, and certain ones were chosen for detailed analysis of prefix, suffix, and stem. A stem-word analyzed in this manner became the nucleus for grouping together other closely related words more or less familiar to the student. The word "traction," encountered in an article on the "Lincoln Highway," brought out a discussion of traction engines, their use in plowing, road-building, and trench warfare, why so called, etc. This centered attention upon the stem "tract." As its meaning became clear the following list was elaborated:

subtract
contract
detract
attract
distract
extract
retract
contraction
attraction
distraction
subtraction
extraction

A study of the prefixes in these words gave a point of leverage for attacking the meaning of words containing them. In this type of prefix study only those words were listed whose stems were familiar to the pupil; as, for example:

recall
reclaim
rearrange
regain
rebound
retain
reform
remake
retake
reinforce
return
reframe, etc.

In a similar manner an acquaintance was made with the most common suffixes.

The meaning of some words was approached by the study of synonyms and equivalent idiomatic phrases. These were, as far as possible, studied in the context and discussed at length to bring out shades of difference in meaning. "An indomitable hero," met in the pioneer tales, brought forth the following synonyms and idiomatic phrases:

indomitable
brave
courageous
resolute
manly
fearless
heroic
bold
daring
plucky
stout-hearted
intrepid
audacious
defiant
undismayed

to look danger in the face
to screw one's courage to the sticking-point
to take the bull by the horns
to beard the lion in his den
to put on a bold front

This type of intensive word-study was continued throughout the first six weeks, but was supplemented by incidental word-study during the remaining twelve weeks.

Oral reading was given special attention during the second six weeks and continued during the following six weeks. The literature was of the same general type as that used in silent reading. The purpose was to improve, if possible, enunciation and expression. Special drills in the enunciation of vowels and of the terminal and initial consonants were a part of each reading lesson. Many of these drills were taken from reading books. Selections were studied silently before being read aloud and the meaning discussed. The various thought units were marked off and the whole selection was then read aloud. Before the close of each lesson the pupil read a selection at sight, unaided by this kind of preparation.

The result. A test in oral reading after this special training showed a reduction of fifty per cent in the number of errors and a gain in rate of reading. In silent reading a greater gain was made, especially in comprehension.

An illustration of group and individual instruction. A fourth-grade teacher[1] has written of her experience in teaching reading to a class of twenty pupils by group and individual instruction based upon the results of testing. Her report is so suggestive that we quote at some length:

> In October pupils were tested to ascertain the oral- and silent-reading rate of each individual. Five oral and five silent trials were made, and the averages obtained and used as measures of reading rate. . . . With but one exception the rapid readers made fewer mistakes. Comprehension was tested informally. Rapidity and comprehension seemed to go together. Intensive instruction was given. Especial attention was paid to poor readers. After two weeks there was no improvement in the rate of the three poorest readers. The only noticeable improvement was made by the better readers. It was evident that the least capable were getting the least from instruction, though receiving more attention. This presented a problem . . . .
>
> Ten types of instruction were planned to cover as many individual needs. The class *Reader* was supplemented by a carefully selected list of books for extensive reading. Methods were devised whereby maximum effort would be called forth and interest sustained. Rate was found to be a measure of improvement which the

---

[1] Zirbes, Laura, "Diagnostic Measurement as a Basis for Procedure"; in *Elementary School Journal* (March, 1918), vol. 18, pp. 505–22.

children could comprehend. They were, therefore, made aware of their rate of reading and kept graphic records of their individual standings in monthly regrouping tests. "A" readers were those whose rate was more than thirteen lines per minute. They were given the privilege of selecting their own material from the supplementary bookshelf for silent reading. This shelf was called "Story Row." The books were arranged in groups according to content. A regular library system was used so that the teacher could ascertain at any time what each child was reading and what he had finished. The quality of the silent reading could thus be revealed by conversation with the pupil. Children who had enjoyed a book were asked to review it for others who might care to read it. Favorite chapters were illustrated. Some children chose informational material. They would recount interesting things which they had learned from their reading, and create a great demand for the book which they had read. No more than two books could be used by a pupil at one time and stories had to be finished before another story book could be begun.

"B" readers were those whose rate was more than nine lines per minute, but not more than thirteen. Pamphlets were provided for their supplementary reading. The material was easier and the content quite suited to their comprehension. Otherwise the system used for the "B" readers was like that for the "A" groups. They had less time for supplementary reading as they required more intensive work with the teacher. Their pamphlets were very popular and were often read by "A" readers.

There was also a group called "C" readers whose rate was between six and nine lines per minute, and another group of "D" readers who read even more slowly and got practically no meaning from the subject-matter. Their supplementary material consisted of separate stories. These they read with the teacher, alternating with her. They liked to have stories read to them. The teacher used her book. The group looked over her shoulder and kept the place, picking up the story and reading on when she stopped, until the end of a paragraph was reached. The meaning was then discussed and the reading continued.

Each child in the class subscribed to a little nature magazine which was kept in the desks for reading during odd minutes. Several other magazines, an atlas, and a file containing good original stories by the children were also at their disposal for this purpose.

The regular reading instruction was the visible means by the aid of which each pupil hoped to get into a group higher than his own by the next measurement.

These groupings were based on rate and were not identical with those made for corrective teaching. The procedures just described, together with the intensive teaching in type lessons which follow, were jointly responsible for improvements in reading rate and quality. This report would, therefore, be incomplete if detailed descriptions of methods used to secure the interest of the individual child were omitted. [Seven of the type lessons relate to oral reading. They are omitted.]

Type lesson 3. Silent reading for the purpose of oral reproduction and comprehension.

Type lesson 4. Silent reading in search of a given phrase, answer, idea, or suggestion in the content of supplementary books, geography text, arithmetic text, and blackboard work.

Type lesson 10. Word-study, with difficult words, for ready recognition, pronunciation, and comprehension. Word-building and word-structure studied.

Results. Fig. 14 shows the results for November, December, and February in silent reading. The range of scores has not been reduced. In fact it has been

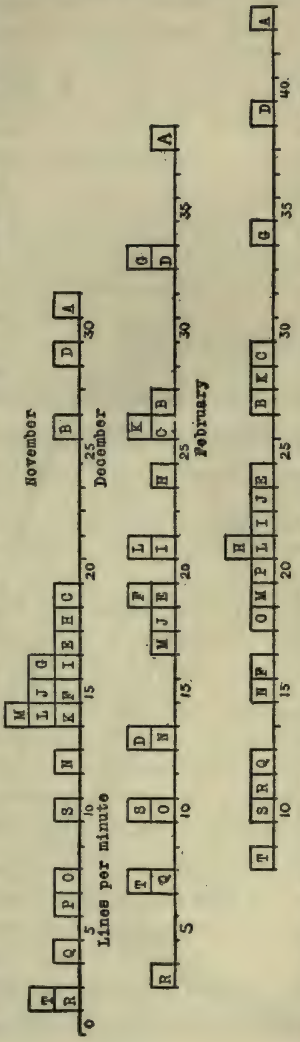


FIG. 14. SILENT READING RATES OF FOURTH-GRADE PUPILS, SHOWING EFFECT OF CORRECTIVE TREATMENT. (*After Zirbes.*)

increased, but the significant thing is that the low scores have been materially increased in most cases. The letters inside the small squares designate the different pupils. Pupil T has advanced from two lines a minute to eight lines a minute. Pupil R has gone from two lines to twelve lines. Pupil S has remained at ten lines. In Fig. 15 we have a graphical representation of one of the causes of this progress; that is, the average amount of supplementary silent reading done during school hours. The increase in the amount is very significant. It shows what can be done when an effort is made.

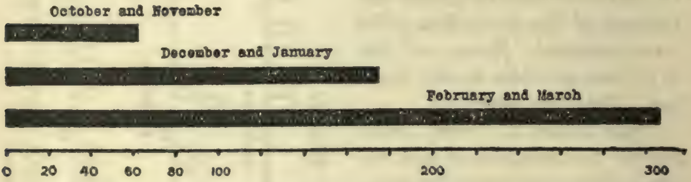


Fig. 15. Average Number of Pages of Silent Reading per Pupil during School Hours. Supplementary Material. (*After Zirbes.*)

Summary for Type III. When the scores of a class are widely distributed, individual instruction is required. Several illustrations of individual defects, the method of dealing with them, and the results have been given. One case of group instruction supplemented by individual instruction has been described.

Type IV. Slow readers and poor in comprehension in primary grades. The class record sheet of a second-grade class which was given the Courtis Silent Reading Test No. 2 is shown in Fig. 16. It is obvious that the pupils of this class both read very slowly and comprehend little of what they read.

The cause: over-emphasis on oral reading. The commonest of all reasons for this situation, particularly when found in the grades below the sixth, is that the teachers have been placing chief stress upon oral reading. Where children are required to give their attention mainly to the correct pronunciation of words, the correct enunciation of sounds, and the correct inflection of the voice in passing over the several punctuation marks, not much growth in the power to comprehend the meaning in the printed page can be expected. Where the children study their reading lesson with the point of view of being able to respond in this way, they fasten upon themselves the habit of watching for words whose pronunciation they are not sure of, or they form the habit of reproducing the sounds of syllables, thus estab-

Table 1
Rate of Reading

SCORE IN WORDS PER MINUTE	NUMBER OF CHILDREN MAKING EACH SCORE.	
Over		
400		
380		
360		
340		
320		
300		
380		
360		
340		
320		
300		
280		
260		
240		
220		
200		
180		
160		
140		
120		
100		
80	3	
60	1	
40	5	
20	5	
0	1	
Total	15	
Median	45	

Index of Comprehension.

Diagnosis		Guesswork			Comprehension poor; additional training needed					Comprehension satisfactory		
Questions Answered	Total	Less than -5	-5 to +5	6-39	40-69	70-79	80-84	85-89	90-94	95-99	100	
70												
65												
60												
55												
50												
45												
40												
35												
30												
25												
20												
15	2	2										
10	5	1		1	3							
5	8	3		1	2	1					1	
0												
Total	15	6		2	5	1					1	

Rate of work satisfactory

Diagnosis: Rate of Reading too slow. Additional training needed.

Median Number of Last Question Answered 7 Median Index of Comprehension 34

FIG. 16. SHOWING THE SCORES OF A 2A CLASS ON THE COURTIS SILENT READING TEST NO. 2. (*Type IV.*)

lishing the practice of moving the lips and other speech organs when reading silently. Frequently both these habits fix themselves upon children whose reading is judged mainly by the daily oral performance. When either or both habits become fixed, a real struggle is required to break them. Unless they are broken, however, the child suffers a severe handicap the rest of his reading life. Many men and women of mature years are still paying the price of those habits fixed in youth. They are able to read but little faster silently than they can pronounce the words orally, because their speech organs make all the motions of the successive words as the reading proceeds.

An illustration of the result of over-emphasis upon oral reading. Judd [1] gives the case of a girl in high school which illustrates the result of over-emphasis upon oral reading. The girl was getting on well in her school work, but "found it exceedingly difficult to keep up with her classes in home assignments." Reports from her various instructors brought out the following statements:

> She is a very satisfactory student in French because she thinks clearly, studies thoroughly, and pronounces easily and correctly. The only drawback to her work is a lack of confidence in herself, which leads her to lose her head occasionally and feel that she knows much less than is the case. In English she is an appreciative and careful student, a little slow at times in getting a grasp of things. She has certainly no serious weakness up to this point and frequently offers hints of superior work. In mathematics she is in the better section and stands eighth among eighty-five students. In general science her work has been very satisfactory and her grades are high.
>
> This girl is like many another student who is getting on all right so far as the school is concerned, but is doing it at great expenditure of effort.
>
> She was tested with a series of passages both in oral and silent

---

[1] Judd, C. H., *Reading: Its Nature and Development.* Supplementary Educational Monographs (University of Chicago Press), p. 161.

reading. . . . The figures show that in general the rates of silent and oral reading are very much alike. . . .

There were marked tendencies to whisper all material read. She was much surprised when told not to do this and was sure she would not understand what she read because, as she said, she understood what she read only when she "heard" the words.

Table 1
Rate of Reading

SCORE IN WORDS PER MINUTE	NUMBER OF CHILDREN MAKING EACH SCORE.	
Over		
400		
380		
360		
340		
320	1	
300		
380		
360		
340		
320		
300		
280		
260	1	
240		
220	1	
200	4	
180	7	
160	4	
140	3	
120	6	
100	1	
80	2	
60	1	
40		
20		
0		
Total	31	
Median	176	

Index of Comprehension.

Diagnosis		Guesswork			Comprehension poor; additional training needed					Comprehension satisfactory		
Questions Answered	Total	Less than -5	-5 to +5	6-39	40-69	70-79	80-84	85-89	90-94	95-99	100	
70												
65												
60												
55												
50												
45												
40	6							1	1	3	1	
35	4						1			1	2	
30	2						1	1				
25	4					1	1		1		1	
20	11						2	4	1	3	1	
15	4					1		1	2			
10												
5												
0												
Total	31					2	5	7	5	7	5	

Rate of work satisfactory

Diagnosis: Rate of Reading too slow. Additional training needed.

Median Number of Last Question Answered 25 Median Index of Comprehension 91

FIG. 17. SHOWING THE SCORES OF A SIXTH-GRADE CLASS ON THE COURTIS SILENT READING TEST NO. 2. (*Type V.*)

The Remedy. The remedy for this type of situation, particularly below the sixth grade, is to place more emphasis upon silent reading. In doing this the suggestions given on page 59 will be helpful.

Type V. Slow readers with a satisfactory index of comprehension. Fig. 17 shows the record of a sixth-grade class which is deficient in rate of work. The median number of words per minute is fifteen less than the standard and the median number of the last question is also fifteen below the standard. The index of comprehension is only four points below standard. These facts show that the pupils of this class read slowly and particularly answer questions slowly. However, they are careful readers as is shown by the index of comprehension. They need training in more rapid reading. The suggestions given on pages 66 and 68, apply to this type also.

VI. Correcting Defects in Oral Reading

The record of the scores of a class in the case of Gray's Oral Reading Test is not so helpful to a teacher as the records of the reading of individual pupils. This shows the types of errors which they made and a knowledge of them frequently will suggest the appropriate remedy.

An illustration of individual instruction in oral reading. The fourth-grade teacher from whose report we quoted on pages 79 to 81 also dealt with oral reading. The "group" instruction described on page 80 applied in part to oral reading. This was supplemented by seven types of individual instruction.

Type lesson 1. All look at the first phrase, looking up when they reach a comma or a period. When the entire group is looking at the teacher she nods and they repeat the phrase. She watches individuals to find their difficulties, but does not interrupt. When they have said all but the last word of the phrase they again look down, silently getting the next phrase and looking up, holding the phrase in mind until all are ready. Again the teacher nods and the group gives the phrase orally, looking down at the last word and continuing this procedure to the end of the paragraph or section.

The intensive study calculated to improve poor readers can be made to yield a double return, if, instead of selecting hard words and subjecting them to analytic study, the unit is the phrase or group of words which expresses an idea. Instead of working at a difficult word, the phrase in which it appears is studied until mastered. Instead of working with one child at a time and giving each child only a few minutes of actual oral reading, four or five of those who have similar ability are grouped together, while other groups of poor readers follow silently. Third-grade material or very simple fourth-grade material is used for this purpose.

While other pupils are being tested, the ones who have had Type 1 answer mentally or in writing blackboard questions concerning the material of their lesson. Occasionally duplicated sheets containing uncompleted sentences or a story are used instead, the children filling in the blanks mentally or in writing.

Type lesson 2. Eye training and focus. Field of vision enlarged to include several words rather than one. First, by having the book far enough from the eyes. Secondly, by eliminating the use of a finger or other place-keeping device. Thirdly, by work with flash cards, flashing phrases, trying to get a phrase at one flash, and counting the number of flashes needed for each phrase. These phrases were cut from current printed matter and mounted on small cards. Written sentences directing children to perform certain activities were also used as flash material. The one who first read the direction carried it out. The one who had three such opportunities in succession was given a sheet with similar work for silent reading and could return to the group when finished.

Type lesson 5. Differentiation for pupils who confuse similar words or miscall syllables, guess at words, or omit endings. Lists like the following form the basis of such work. Lists are compiled from actual mistakes made by children:

that
when
what
then
how
who
their
there
than
women
man

woman
every
never
even
ever
very
these
those
now
know
beautiful

beautifully
prettiest
prettily
probably
lovingly
companions
understand
understood
laughingly
quietly
left

swimming
board
close
chose
lying
buying
tired
tried
certain
curtain
felt

Type lesson 6. Lessons in accuracy for those who make errors, substitutions, and omissions; reading a page orally and counting errors, or reading until they make an error to see how many lines they can read perfectly.

Type lesson 7. Breathing exercises. Children taught to breathe rhythmically at ends of phrases or clauses instead of breaking the smoothness of oral reading. Practice in breath control is thus related to the problem of meaning and interpretation. Abdominal breathing taught.

Type lesson 8. Articulation exercises for mumblers, or those with other bad speech habits.

Type lesson 9. Voice work and expression. Unpleasant voice quality and monotony corrected by special practice and training. Children are taught to vary meaning by change of stress and to show relative importance of ideas similarly. Punctuation is studied for the same purpose.

The Results. For oral reading very striking results were obtained. They are shown graphically in Fig. 18. The increase in the closeness of the grouping of the members of the class is significant as well as the progress made by the individual pupils.

The base line represents lines read per minute. The lettered blocks represent individuals. The group divisions are shown by vertical lines. Thus pupil O was in Group "D" and read five lines per minute in November, moved to Group "C" in December, and was reading nine lines per minute. In January he read twelve lines per minute and was in Group "B." In February he read thirteen lines per minute, but did not get into Group "A" until March. . . . The reading rate is an average of from three to five trials on as many selections of unstudied material from the Horace Mann *Reader*. No rate of measurement was made in April, as other reading tests were conducted.[1]

Interpretation of scores in ungraded schools. In rural schools or other ungraded schools classes are frequently so small that it is not wise to use the median or class scores. The interpretation is largely a matter of comparing the scores of individual pupils with the standards. In doing this certain facts should be kept in mind. First, the standards are median scores. For example, the standards for the fifth

---

[1] *Elementary School Journal,* vol. 18, p. 512.

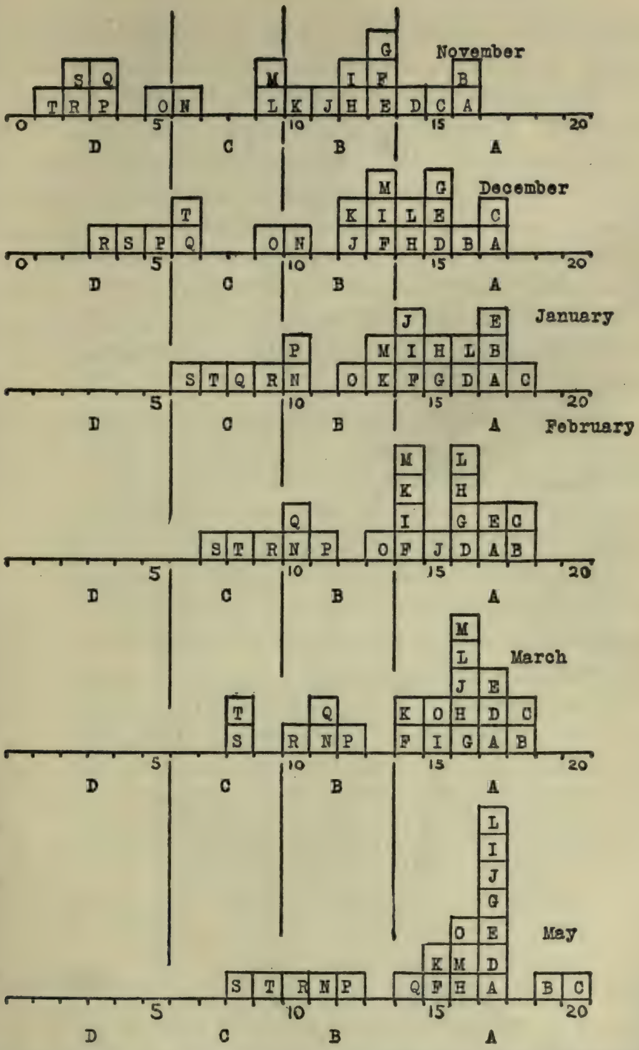


FIG. 18. SHOWING IMPROVEMENT IN ORAL-READING RATE OF TWENTY-FOURTH WHEN INDIVIDUAL AND GROUP INSTRUCTION WAS USED. RATE EXPRESSED IN LINES PER MINUTE. (*After Zirbes.*)

grade are the median scores of several thousand fifth-grade pupils. Thus half of these pupils had scores greater than the standard scores and half of them had scores less than the standards. Second, pupils of the same grade differ widely in ability, and when the class scores are up to standard, half of the class will be above standard and half will be below. Third, in schools where only a few pupils belong to a grade, it may be that all are pupils of little natural ability, and hence should not be expected to have scores up to standard. On the other hand, it may be that they have a high degree of natural ability and should have scores distinctly above standard.

This makes the interpretation of the scores of individual pupils or of small groups of pupils difficult and somewhat uncertain. However, it is safe to say that when a pupil is distinctly below standard in rate or comprehension of silent reading, or both, his case should be carefully studied by the teacher. Much of what has been said concerning the causes of class weaknesses and the correctives for them also applies to individual pupils. Some pupils will be found to be high in rate, but low in comprehension (the rapid, careless reader; see pages 51 to 58); others will be low in rate, but relatively high in comprehension (the slow, plodding, but careful reader; see pages 65 to 68); still others will be low in both rate and comprehension (the slow, careless reader or one who has not learned to read; see pages 73 to 79).

Some pupils will be found who are distinctly above standard in both rate and comprehension. In such cases the teacher's problem is different. These have made satisfactory progress in silent reading under the instruction which they have been receiving and thus they require no special attention. The only case in which a modification of the instruction should be made is when a pupil who is above standard in silent reading is not doing satisfactory work in arithmetic, spelling, or some other subject. Then he should spend less time reading and more upon the subject in which he is weak.

The rural teacher's opportunity for diagnosis and corrective instruction. When a teacher has a small number of pupils, he has an unusual opportunity for diagnosing his pupils and applying the required corrective instruction. Each pupil can be studied until his strong points and his weaknesses are known. Upon the basis of this information the teacher can easily apply the correctives because the instruction is largely individual anyway. The teacher who has twenty-five or more pupils in one class has much less opportunity to deal with individual pupils. When the teacher knows the scores of the pupils of his class on standardized tests, each pupil can be given that kind of instruction which he needs and under which he will make the greatest progress of which he is capable.

The use of standardized tests does not require a large amount of time. To teachers who may be interested in such diagnostic and remedial work this comment by the author of the above report is significant. (See pp. 79 and 86.)

> This study was not conducted by a specialist, but by a grade teacher interested in the advancement of the class through methods which reach the individual members. No time was taken from other studies. In fact a similar experiment in individual instruction was simultaneously carried on in spelling, arithmetic, and penmanship. The results were tested and are evidence that no one subject was over-stressed. The time economy, resulting from scientific procedure, also made possible a fullness and breadth of teaching usually thought incompatible with standardization and educational measurement.[1]

Summary. The ideas presented in this chapter may be summarized under three heads:

(*1*) *Service of standardized tests to the teacher.* By far the largest service of standardized tests in reading is being rendered to the teacher. Not only are they enabling the teacher to check up his conception of what can justly be expected of children, but they are indelibly impressing upon his mind the absolute need for recognizing the individual differences among his pupils in respect to each problem of learning, and for studying the reading needs of his pupils in order to plan the instruction most wisely. In this chapter we have considered the meaning of certain types of class records, how to secure additional information concerning the reading abilities of pupils and the general correctives which should be applied to improve the standing of such classes. For certain cases we have given illustrations of the detailed diagnosis of pupils, the correctives which were applied, and the results which were obtained.

---

[1] *Elementary School Journal* (March, 1918), p. 522.

The distinction between silent reading and oral reading has been emphasized. The teacher should bear in mind that a pupil may read well orally from his reader, but may be doing poorly in geography or in the problems of arithmetic because he has not learned to read silently. These content subjects depend upon reading, but not upon the sort of oral word-pronouncing which still too largely characterizes our reading periods. Such a child needs a different sort of reading. He would be found to stand low, probably, in vocabulary; probably low in quality of silent reading. If the teacher has before him a chart upon which is recorded the standing of this child in the various aspects of reading, he will no longer assign for his study the next page or two in the *Reader*. He needs the sort of reading which widens his vocabulary more rapidly and centers his thought upon meaning instead of upon words.

As an illustration [1] of the aid of reading tests in such diagnosis, the case of the Training School at Oshkosh, Wisconsin, may be cited. During a summer term of only six weeks, pupils, by use of the Kansas Silent Reading Tests, the Gray Oral Test, and the Gray Silent Reading Tests, had their difficulties localized. Instruction was then given upon the points revealed to be needing attention. Twenty out of one hundred and five children were given different instruction from that given the class as a whole. Surprisingly greater results were obtained in the case of those children whose instruction was specifically adapted to their difficulties.

---

[1] Uhl, W. L., "The Use of the Results of Reading Tests as a Basis for Planning Remedial Work"; in *Elementary School Journal* (December, 1916), vol. 17, no. 4.

(*2*) *Service of standardized tests to the rural teacher.* Standardized tests render a peculiar service to rural teachers by setting standards for them. A teacher who works in isolation from other teachers needs to know how her pupils compare with pupils in other schools. Since standardized tests have been widely used, they furnish a means by which any teacher can easily ascertain how her pupils stand in comparison with the established standards.

(*3*) *Service of standardized tests to the child.* Since the beginning of schools children have been sent to school to be taught. That being the case, they wait to be told what to do, and there is the end of their responsibility. When the end of the month comes they look to their report card for a measure of their success in doing what they have been told.

A function of standardized tests, by which the child can measure his own achievements about as successfully as the teacher can, is that they bring the child into partnership with the teacher in directing the whole educative process. If the child discovers by actual trial that he has only three fourths as large a vocabulary as children of his grade the country over, or that he reads only three fourths as fast, he can be depended upon better to coöperate in overcoming the fault than when he is simply given a card every month with 70 assigned to his reading. Particularly is this true if he feels that at the end of a given period he can take his own measure again to ascertain his gain. Children should be enlisted with the teacher in the effort to select the most needful sorts of materials for their study. Where one child needs problem-solving, another needs a story, while still another needs something else than reading of any kind.

Service of standardized tests to the superintendent. Although this book is written primarily for teachers, it will not be out of place to call attention to the usefulness of standardized tests to the superintendent or supervisor. From the importance of reading in the general efficiency of all school work we may assume that the superintendent is vitally interested in making the instruction in reading most effective. What can reading tests reveal to him?

First, they can satisfy him and his teachers of the general status of reading in his district. It is easy for any superintendent to carry conviction among his teachers that the results in reading are not satisfactory in his district if he can show that among a group of a dozen or more neighboring cities his district stands low. The extent to which it stands low becomes a measure of the renewed earnestness needed in attacking the problem of improvement.

It is difficult for one to carry in mind a fixed standard of achievement. One gradually thinks more and more in terms of what those around him are achieving. It would have been quite impossible, for example, to convince the superintendent and teachers of the Anglo-Korean school at Songdo,[1] without a standardized test, that the children in their fifth grade could do, on the average, reading work valued at only 3.8 units, while American children, who had been in school only the same number of years, could score 13.2 units, or that their sixth grade could accomplish only as much as the American third grade. It meant much to that school for its superintendent and teachers to be able to measure their school by the American standards.

---

[1] Wasson, Alfred W., "Report of an Experiment in the Use of the Kansas Silent Reading Tests with Korean students"; in *Educational Administration and Supervision*, vol. 3, p. 98.

Reveal wrong emphasis in teaching. Differences in the reading work done in the several buildings within a city may be as striking as differences among cities. In a certain Middle-Western town a forceful principal of one of the ward buildings has dominated the work of the building for a good many years. The reading of the building was his particular pride. When tested for silent reading ability his children scored in every grade but little more than half what the children in another building scored where the work was reputed to be "much less thorough." These results were made the basis of deliberations among the teachers as to the legitimate outcomes of reading, with the result that, without diminishing any one's zeal, the emphasis was transferred from oral word-pronouncing to silent thought-getting in the building where this strong principal dominates the work so effectually.

QUESTIONS AND TOPICS FOR STUDY

1. What are the chief methods by which adults add new words to their vocabularies? Are more new words learned from the context in which they appear, or from the dictionary? What can you say concerning the best way to increase the vocabulary of children?
2. What are some of the other factors besides vocabulary involved in silent reading? In what grades is vocabulary the most important factor? Make some suggestions for guaranteeing the intimate association of the mental concept which a word symbolizes, and the word itself when it is encountered in word drills.
3. What is the significance of rate in reading? Is there any truth in the rather common belief that one who reads slowly "gets more out of what he reads"? If you do not know the answer, can you devise some

way to test it out in your class? Compare your own silent reading rate with that of some equally well-educated friends.

4. What are the chief dangers involved in having much oral reading in the lower grades? Can these dangers be safeguarded? What types of reading matter do you now read orally outside the schoolroom? Are these the types which your pupils are asked to read orally?
5. What are the circumstances under which you last read aloud? Do your pupils have the same incentives for reading clearly and interestingly that you had on that occasion?
6. What are some of the things you do to assist your pupils in developing ability to comprehend the meaning of the printed page? Do you know of faulty habits which some of them have which prevent their centering attention upon the meaning? Do you know which pupils read with accuracy? Which with rapidity?
7. How long does it take you to become familiar with the reading difficulties of each child when you receive a new class of, say, thirty children? Would you consider it economical if some tests were available by means of which you could discover these difficulties as well as others the first day and thus prepare a chart of each child's instructional needs? How long at the beginning of a term could you afford to spend in making such a diagnosis?
8. Think of the last examination you gave in reading. Did it test satisfactorily what you are striving to teach in reading?

CHAPTER IV

THE MEASUREMENT OF ABILITY IN THE OPERATIONS OF ARITHMETIC

How arithmetical ability is measured. The plan of measuring the ability of pupils to do the operations of arithmetic is to have them do a set of examples[1] under specified conditions. In order that the scores may have a definite meaning, the test is limited to one type of example in one operation, as subtraction or multiplication or at most to a group of closely related types. The necessity for determining the value of the different examples in the test can be eliminated by constructing the examples so that they contain the same number of combinations; that is, in addition having each example involve twenty additions. The rate at which the pupil performs the operations is measured by timing the test as was done in the measurement of the ability to read silently. In the following pages we will describe two tests which have been devised to measure the ability of pupils to do the operations of arithmetic.

I. Courtis Standard Research Tests, Series B

Description of tests. The Standard Research Tests, Series B, or, as they are commonly called, the Courtis Arithmetic Tests, have been more widely used than any other instrument for measuring arithmetical abilities, and as a result we have better comparative standards for their use. The series consists of four tests, printed on four consecutive pages. They are suitable for a general survey of the abilities of pupils to perform the operations with integers. They are used in Grades four to eight.

---

[1] The word "problem" is used by some writers to designate both "example" and "problem." In this book the word "example" will be used to designate exercises which explicitly call for certain arithmetical operations. The word "problem" will designate only those exercises which require the pupil to determine first what operations are to be performed.

Test No. 1. Addition

The twenty-four examples of this test have been constructed so that all have the same form, three columns of nine figures each. The following are samples. Time allowed, 8 minutes.

927	297	136	486	384	176
379	925	340	765	477	783
756	473	988	524	881	697
837	983	386	140	266	200
924	315	353	812	679	366
110	661	904	466	241	851
854	794	547	355	796	535
965	177	192	834	850	323
344	124	439	567	733	229

In giving the test the pupils are directed as follows:

You will be given eight minutes to find the answers to as many of these addition examples as possible. Write the answers on this paper directly underneath the examples. You are not expected to be able to do them all. You will be marked for both speed and accuracy, but it is more important to have your answers right than to try a great many examples.

Test No. 2. Subtraction

This test consists of twenty-four examples, each involving the same number of subtractions. The following are samples. Time allowed, 4 minutes.

107795491	75088824	91500053	87939983
77197029	57406394	19901563	72207316

Test No. 3. Multiplication

This test consists of twenty-four examples of this type. Time allowed, 6 minutes.

$$\begin{array}{r}8246\\ \underline{29}\end{array}\qquad\begin{array}{r}3597\\ \underline{73}\end{array}\qquad\begin{array}{r}5739\\ \underline{85}\end{array}\qquad\begin{array}{r}2648\\ \underline{46}\end{array}\qquad\begin{array}{r}9537\\ \underline{92}\end{array}$$

Test No. 4. Division

This test consists of twenty-four examples of this type. Time allowed, 8 minutes.

$$25)\underline{6775}\qquad 94)\underline{85352}\qquad 37)\underline{9990}\qquad 86)\underline{80066}$$

$$73)\underline{58765}\qquad 49)\underline{31409}\qquad 68)\underline{43520}\qquad 52)\underline{44252}$$

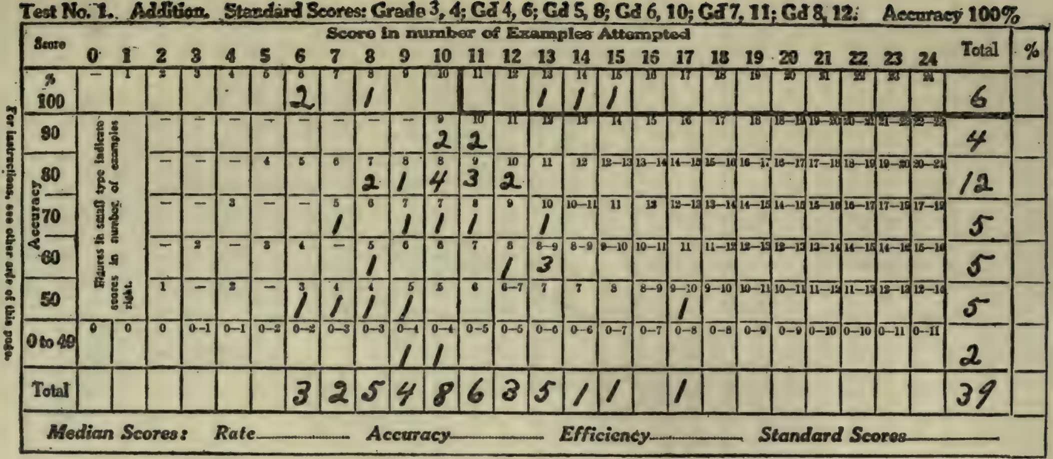
Each of the examples of a test calls for the same number of operations under approximately the same conditions. This makes the examples of each test approximately equal in difficulty. Any example of the addition test, say the seventh, is just as difficult as any other, say the second. Thus, the tests consist of twenty-four equal units, just as a yardstick consists of thirty-six equal units (inches). The measure of a pupil's ability is represented by the distance he advances along the scale in the given time; that is, by the number of examples done and by the per cent of these examples which have been done correctly.

Since an example of one of these tests is defined as so many operations under certain conditions, it is possible to construct other tests equal in difficulty. Four forms have been constructed. This makes it possible to use a different form when the tests are given a second time.

Giving the tests. The general directions for giving reading tests also apply in the case of arithmetic. (See page 25.) Detailed directions accompany these tests. Since the same tests are given in all grades, a group of pupils belonging to two or more grades can be tested together. It is only necessary to sort the papers according to classes when recording scores. If it is not convenient or desirable to give the four tests on one day, the test papers may be collected and returned the next.

Marking the papers. In marking the test papers, which is done by the use of a printed answer card that is run along across the page, no credit is given for examples partly right or for examples partly completed. This simple plan of marking the papers insures uniformity and accuracy. A pupil's score is the number of examples attempted and the number right. The number of examples attempted is a measure of the pupil's *rate of work*. By dividing the number right by the number attempted, the per cent of examples correct may be obtained. This is a measure of the *quality* or *accuracy* of his work. Thus, the two "dimensions" of the ability to do the operations of arithmetic are *rate* and *accuracy*.

Recording the scores of a class. For recording the scores of a class a record sheet of the form shown in Fig. 19 is used. This figure contains merely the blank for addition, but the forms for the other three tests of the series are identical with it. Detailed instructions for recording scores are printed on the record sheet. The large figures at the top of the sheet refer to the number of examples attempted and the small figures within the squares refer to the number of examples done correctly. The sheet is arranged so that the per cent of examples done correctly is computed automatically and the distribution of the scores according to both rate and accuracy is obtained at the same time. The scores of a seventh-grade class are shown in Fig. 19. The numbers written in certain of the squares represent the number of pupils whose scores fell within these divisions of the record sheet. The distribution according to rate is found at the bottom of the record sheet and is to be read thus: Three



Test No. 1. Addition. Standard Scores: Grade 3, 4; Gd 4, 6; Gd 5, 8; Gd 6, 10; Gd 7, 11; Gd 8, 12. Accuracy 100%

	Score	Score in number of Examples Attempted																									Total	%
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
	% 100	—	1	2	3	4	5	6 2	7	8 1	9	10	11	12	13 1	14 1	15 1	16	17	18	19	20	21	22	23	24	6	
	90	Figures in small type indicate scores in number of examples right.		—	—	—	—	—	—	—	—	9 2	10 2	11	12	13	14	15	16	17	18	18–19	19–20	20–21	21–22	22–23	4	
Accuracy	80			—	—	—	4	5	6	7 2	8 1	8 4	9 3	10 2	11	12	12–13	13–14	14–15	15–16	16–17	16–17	17–18	18–19	19–20	20–21	12	
	70			—	—	3	—	—	5 1	6	7 1	7 1	8 1	9	10 1	10–11	11	12	12–13	13–14	14–15	14–15	15–16	16–17	17–18	17–18	5	
	60			—	2	—	3	4	—	5 1	6	6	7	8 1	8–9 3	8–9	9–10	10–11	11	11–12	12–13	12–13	13–14	14–15	14–15	15–16	5	
	50			1	—	2	—	3 1	4 1	4 1	5 1	5	6	6–7	7	7	8	8–9	9–10 1	9–10	10–11	10–11	11–12	11–13	12–13	12–14	5	
	0 to 49	0	0	0	0–1	0–1	0–2	0–2	0–3	0–3	0–4 1	0–4 1	0–5	0–5	0–6	0–6	0–7	0–7	0–8	0–8	0–9	0–9	0–10	0–10	0–11	0–11	2	
	Total							3	2	5	4	8	6	3	5	1	1		1								39	

Median Scores: *Rate*________ *Accuracy*________ *Efficiency*________ *Standard Scores*________

For instructions, see other side of this page.

FIG. 19. SHOWING FORM OF TABULATION SHEET FOR RECORDING SCORES OBTAINED BY USING THE COURTIS STANDARD RESEARCH TESTS, SERIES B, AND THE SCORES OF A SEVENTH-GRADE CLASS IN ADDITION.

pupils attempted only six examples, two pupils attempted only seven examples, five pupils attempted only eight examples, etc. The distribution according to accuracy is found at the right-hand side of the sheet and is to be read thus: The per cent of examples done correctly by two pupils was less than fifty per cent, for five pupils it was between fifty per cent and sixty per cent, for five other pupils it was between sixty per cent and seventy per cent, etc. These two distributions, the one according to the per cent of examples done correctly and the other according to the number of the examples attempted, describe the ability of the pupils of this class to do the examples of the addition test.

Calculating the median score of a class. The central tendency (median) of the distribution of scores in the case of silent reading was found by arranging the test papers in order and counting up to the middle paper. The score on this paper was the median score. (See page 28.) This is not the best method to use here. For calculating the median, Courtis gives explicit directions in Folder D which is designed to accompany this series of tests. We give here general directions for finding the median so that they can be conveniently referred to.

The median is the mid-measure of a distribution. In the case of a distribution having an odd number of scores it is the value of the middle score. If there is an even number of scores it is halfway between the two middle scores.

Table XII gives a group of scores in the form of a distribution. The column labeled "Score" gives the scores arranged in order of magnitude. The column headed "Frequency" tells how many scores there are of each magnitude, or how frequently each score occurs. In this table there are two scores of 13, two scores of 12, five scores of 11, etc. The total number of scores is 45. The twenty-third score is the middle one, but it is not possible to identify its value di-

TABLE XII. SHOWING A TYPICAL DISTRIBUTION OF SCORES IN NUMBER OF EXAMPLES ATTEMPTED

Score	*Frequency* *
14	1
13	2
12	2
11	5
10	9
9	7
8	8
7	5
6	3
5	2
4	1
3	
2	
Total	45
Approximate median	9.0
Correction	.6
True median	9.6

* The frequency is the number of pupils making the score indicated. The scale consists of the scores arranged in order.

rectly. It is clearly one of the seven scores given as 9, because counting from the lower end of the distribution, 1 and 2 are 3, and 3 are 6, and 5 are 11, and 8 are 19, and 7 are 26, which is beyond the middle of the distribution. Therefore the approximate value of the median of this distribution is 9.0, which is the interval which contains the true median.

Rule for finding median. (1) Find the half sum of the distribution. In case there is an even number of scores, there is no middle score. In such a case the average of the two most central scores may be taken, although for practical purposes it will be satisfactory to take the lesser of the two middle scores. By doing this the calculation of the median is made simpler. Thus, if a distribution contains forty-one scores, the middle score is the twenty-first score; if it contains forty scores, the twentieth score may be taken as the middle score.[1] The number of the middle score is obtained by dividing the total number of scores by two. This quotient, expressed as the nearest integer, is called the "half sum." When using a particular test the directions which accompany that test should be followed, because the medians used for standards have been obtained by following the accompanying directions.

(2) To determine the approximate median it is simply necessary to locate the interval of the distribution in which the median falls. To locate the interval in which the median falls, begin at the lower end of the distribution and add together the frequencies until the addition of the next one will make a sum greater than the number of the median score, or half of the total of the frequencies. This sum of the frequencies is called the "partial sum." The median score is in the next interval, and the approximate median is the value of that interval.

(3) To calculate the amount to be added to the approximate median to make the true median, proceed as follows: (1) Subtract the partial sum of the frequencies from the half sum. The partial sum is found in determining the approximate median. (2) Divide this difference by the number of scores which are included in the interval in which the true median falls. Add this quotient to the approximate median.

---

[1] There is a difference of practice on this point. The directions which accompany the Monroe Standardized Silent Reading Tests read as follows: "The median score is the score on the middle paper in the pile of papers arranged according to size of scores. If there are thirty-five papers, the median score is the score on the eighteenth paper. If there are thirty-six papers, the median score is halfway between the score on the eighteenth paper and the score on the nineteenth paper."

The directions which accompany the Courtis Standard Research Tests in Arithmetic, Series B, read as follows: "If there are thirty-seven children in the class, the nineteenth score in order of magnitude would be the median score; for there would be eighteen scores larger and eighteen smaller. If there were thirty-six children in the class, the eighteenth score would be taken as representing the nearest approximation to the middle measure."

If the width of the interval is more than one unit, the quotient must be multiplied by the number of units the interval contains. It is well to carry the quotient to two decimal places, but in writing the median it should be expressed only to the nearest tenth.[1]

The rule applied. In Table XII the half sum is 23. The approximate median is 9.0. The partial sum being 19, four of the seven scores in the 9-interval are required in order to reach the mid-point of the distribution. Four divided by 7 is .6, which, added to 9.0, gives 9.6 the true median.

Special cases. Although the rule given applies to all cases, there are a few special cases which sometimes give trouble. In Table XIII we show three special cases which may arise in using Series B. Case A is where the partial sum (13) is also the half sum (13). The approximate median is in the *next* interval (9). Since the difference between the partial sum and the half sum is zero, there is no correction and the true median is also 9.0. Case B is where the median falls in the first interval (0 to 49). It is only necessary to remember that the width of this interval is 50. Case C is where the median falls in the 100-interval. The width is zero. Hence the correction is zero.

Efficiency. Courtis has defined a measure of "efficiency." He says:

> The word "efficiency" as applied to education has, only too frequently, but little meaning. As used in connection with the Courtis Tests, however, it has a very definite meaning as soon as the following definition has been accepted: *The efficiency of any teaching process which has a measurable product is the per cent of the total product that measures up to the standard for the grade.*
>
> For each test find the sum of all the frequencies equal to or exceeding the standard. Multiply this sum by 100 and divide by the total number of scores. The result will be the efficiency for that test.

---

[1] See King, W. I., *The Elements of Statistical Method*, pp. 129–30; and Thorndike, E. L., *Mental and Social Measurements*, p. 54.

In Fig. 19 lines have been drawn to mark off those scores which are up to standard in both rate (11) and accuracy (100). There are three such scores. The efficiency of this class thus is 300÷39, or 7.7. However, many users of Series B do not believe the efficiency score has an important significance and it is not generally used.

TABLE XIII. SHOWING THREE SPECIAL CASES IN CALCULATING THE MEDIAN WHICH ARISE IN USING COURTIS'S STANDARD RESEARCH TESTS, SERIES B

A		*B*		*C*	
Scale	*Frequency*	*Scale*	*Frequency*	*Scale*	*Frequency*
15	..	100	1	100	15
14	..	90– 99	..	90– 99	1
13	1	80– 89	2	80– 89	4
12	1	70– 79	2	70– 79	3
11	2	60– 69	1	60– 69	2
10	3	50– 59	4	50– 59	..
9	5	0– 49	14	0– 49	..
8	4				
7	6	Total	24	Total	25
6	3				
5	..	Approximate median	0.0	Approximate median	100
4	..	Correction	43.0	Correction	0
Total	25	True median	43.0	True median	100
Approximate median	9.0				
Correction	.0				
True median	9.0				

Standard median scores. In Table XIV there are given three standard scores. (1) General median scores based upon distributions of "many thousands of individual scores in tests given in May or June, 1915–16. The distribution for each grade was made up of approximately equal numbers of classes from large-city schools and from small-city and country schools." (2) The standards proposed by Courtis after three years' use of these tests. (3) Boston standard median scores after the tests had been used for three years.

These standards are given in terms of rate and accuracy, which is the best form. However, for certain purposes it may be desirable to have them in terms of "number attempted" and "number right." The number of examples right can be found by multiplying the number of "examples attempted" or the rate by the accuracy.

With reference to the standards which he has proposed Courtis says:

> The speeds [rates] set as standard are approximately the average speeds [rates] at which the children of the different grades have been found to work when tested at the end of the year, when for any one grade a random selection of five thousand scores from children in schools of all types and kinds are used as a basis of judgment.
>
> Standard accuracy is perfect work, one hundred per cent. This is a tentative standard only, as there is available very little information in regard to the factors that determine accuracy and the effects of more efficient training.
>
> At present in addition and multiplication it is only very exceptional work in which the median rises above eighty per cent accuracy, while in subtraction and division the limiting level is ninety per cent.
>
> Standard speeds [rates] are not likely to change greatly. Standard accuracy is surely destined to approach much more nearly one hundred per cent than present work would indicate.
>
> Standard scores are not only goals to be reached; they are limits not to be exceeded. It seems as foolish to overtrain a child as it is to undertrain him. All direct drill work should, in the judgment of the writer, be discontinued once the individual has reached standard levels. If his abilities develop further through incidental training, well and good, but the superintendent who, by repeated raising of standards, forces teachers and pupils to spend each year a larger percentage of time and effort upon the mere mechanical skills, makes as serious a mistake as the superintendent who is too lax in his standards.[1]

---

[1] Courtis, S. A., *Third, Fourth, and Fifth Annual Accountings, 1913–16* (Department of Coöperative Research, Detroit), p. 49.

TABLE XIV. STANDARD MEDIAN SCORES, COURTIS'S STANDARD RESEARCH TESTS, SERIES B

Grade	*Addition*		*Subtraction*		*Multiplication*		*Division*	
	Rate	*Accuracy*	*Rate*	*Accuracy*	*Rate*	*Accuracy*	*Rate*	*Accuracy*
IV. General	7.4	64	7.4	80	6.2	67	4.6	57
Courtis	6	100	7	100	6	100	4	100
Boston	8	70	7	80	6	60	4	60
V. General	8.6	70	9.0	83	7.5	75	6.1	77
Courtis	8	100	9	100	8	100	6	100
Boston	9	70	9	80	7	70	6	70
VI. General	9.8	73	10.3	85	9.1	78	8.2	87
Courtis	10	100	11	100	9	100	8	100
Boston	10	70	10	90	9	80	8	80
VII. General	10.9	75	11.6	86	10.2	80	9.6	90
Courtis	11	100	12	100	10	100	10	100
Boston	11	80	11	90	10	80	10	90
VIII. General	11.6	76	12.9	87	11.5	81	10.7	91
Courtis	12	100	13	100	11	100	11	100
Boston	12	80	12	90	11	80	11	90

Comparisons of class scores with the standards given in Table XIV or any others are valid only when the tests have been given under standard conditions. Slight changes in the method of giving the tests may affect the scores as much as the difference in the standards from one grade to another.

The interpretation of scores. The standards for Series B are to be used for the interpretation of individual and class scores in much the same way as the standards for the reading tests. The form of the class record sheet is convenient for interpreting the scores of the class. If one will draw a vertical line to represent the standard rate and a horizontal line to represent the standard accuracy for the grade, one can tell at a glance what the condition of the class is. See Fig. 21, page 120.

Graphical representation of the scores of a school. In Fig.

20 there is shown a scheme devised by Courtis for the graphical representation of the median scores of a city or school building. The position of the standards is shown by the circles along the dotted-line curve. The position of the median scores of a city is shown by the X's through which the solid-line curve passes. This is a very effective means for showing the condition within a city or school. The figure makes it very clear that the median scores for number of examples attempted are conspicuously below standard. The position of the first X which represents the fourth-grade scores is below and to the left of the fourth-grade circle. This means that the fourth-grade scores are below standard. The fifth grade shows an increase in accuracy but the pupils do not work more rapidly than those in the fourth grade. From the fifth grade to the sixth and from the sixth grade to the seventh growth is shown, principally in accuracy. For number of examples attempted the seventh grade is below the fifth-grade standards. The eighth grade shows an increase in rate of work but a marked decrease in accuracy.

II. MONROE'S DIAGNOSTIC TESTS

Arithmetical abilities distinct. A few years ago Stone[1] investigated the nature of ability in arithmetic and concluded that it was made up of a number of specific abilities. His conclusions have been corroborated by a number of other investigations,[2] and it is now reasonably certain that

---

[1] Stone, C. W., *Arithmetical Abilities and Some Factors Determining Them.* (Teachers College Contributions to Education, no. 19, 1908.)

[2] Kallom, Arthur W., *Determining the Achievement of Pupils in Addition of Fractions.* (School Document, no. 3, 1916. Boston Public Schools.)

Recently an investigation was made, under the direction of the writer, of the nature of the ability to place the decimal point in a quotient. This investigation showed that a number of specific abilities were involved, and not a single ability. See Monroe, Walter S., "The ability to place the decimal point in division," *Elementary School Journal,* vol. 18, pp. 287–93 (December, 1917).

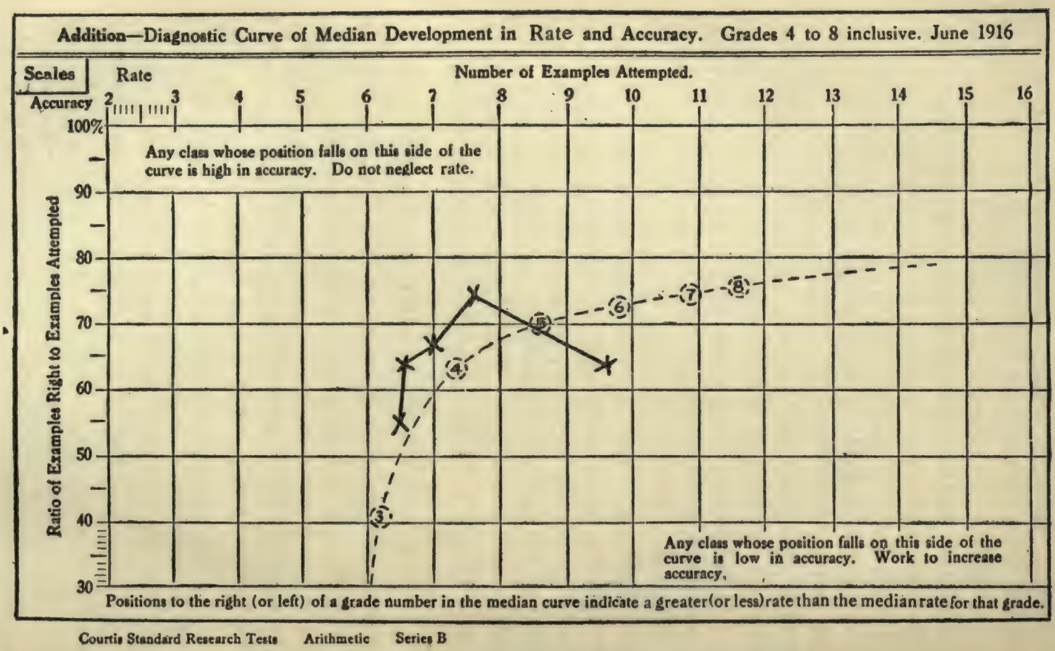


FIG. 20. SHOWING USE OF COURTIS'S GRAPH SHEET NO. 3.

in teaching the operations of arithmetic, we are attempting to engender a number of specific abilities which are relatively distinct, and not a single arithmetical ability. The word "ability" is used to refer to the rate and accuracy with which a pupil does a certain type of example. Teachers have recognized that pupils could do subtraction examples in which there was no "borrowing" when they were unable to do examples in which there was "borrowing," or that they could do short division when they were unable to do long division. The investigations of Stone and others have proven that there are as many different abilities as there are types of examples. In fact, it is obvious that the ability to add a column of three figures is not the same as the ability to add a column of twelve figures. In adding a column of figures it is necessary that one hold in mind the partial sum until he has added the next figure. This process must be repeated until the final sum is reached, and a failure to do this continuously will result in stopping the adding, at least temporarily. It is a frequent occurrence, for one who is not accustomed to adding long columns of figures, to find that he has stopped, perhaps has even lost the partial sum, and must begin again. The span of attention required in adding three figures is short, and pupils who are able to do examples of this type with a high degree of skill frequently are unable to add long columns of figures with an equal degree of skill. In fact, we have no reason to expect them to be able to do this type of example until they have practiced upon it.

Courtis,[1] the author of the Standard Research Tests, has identified the following types of examples in the operations with integers:

Addition: (1) addition combinations; (2) single-column addition of three figures each; (3) "bridging the tens," as 38 + 7; (4) column addition, seven figures; (5) addition with carrying; (6) column addition with increased attention span, thirteen figures to the column; (7) addition of numbers of different lengths.

---

[1] Courtis, S. A., *Teacher's Manual for Courtis Standard Practice Tests* (1916.)

Subtraction: (1) subtraction combinations; (2) subtraction of 9 or less from a number of two digits, without "borrowing"; (3) same as the second, but with "borrowing"; (4) subtraction of numbers of two or more digits involving borrowing.

Multiplication: (1) multiplication combinations; (2) multiplicand two digits, multiplier one digit, and no carrying; (3) same as number 2, but with carrying; (4) long multiplication, without carrying; (5–8) zero difficulties, four types:

$$\begin{array}{r}560\\ \underline{40}\end{array}\qquad\begin{array}{r}807\\ \underline{59}\end{array}\qquad\begin{array}{r}617\\ \underline{508}\end{array}\qquad\begin{array}{r}703\\ \underline{60}\end{array}$$

(9) long multiplication, with carrying.

Division: (1) division combinations; (2) simple division, no carrying; (3) same as number 2, but with carrying; (4) long division, no carrying; (5–6) zero difficulties, two cases:

$$\begin{array}{r}690\\ 71\overline{)48990}\end{array}\qquad\qquad\begin{array}{r}302\\ 31\overline{)9362}\end{array}$$

(7) long division, with carrying, "first case, the first figure of the divisor is the trial divisor and the trial quotient is the true quotient":

$$\begin{array}{r}72\\ 63\overline{)4536}\end{array}$$

(8) "second case, where the trial divisor is one larger than the first figure of the divisor, but the trial quotient is the true quotient":

$$\begin{array}{r}63\\ 49\overline{)3087}\end{array}$$

(9) "third case, where the first figure of the divisor is the trial divisor, but the true quotient is one smaller than the trial quotient":

$$\begin{array}{r} 89 \\ 63\overline{)5607} \end{array}$$

(10) "fourth case, where the first figure of the divisor must be increased by one to obtain a trial divisor and the second trial quotient must be increased by one to get the true quotient)":

$$\begin{array}{r} 79 \\ 36\overline{)2844} \end{array}$$

Each a specific habit. Each of these types of examples requires a specific habit or automatism. To be sure, certain elements, such as the fundamental combinations, are common, but careful analysis will show that the ability to do examples of one type is different from that required to do another. Not only will a careful analysis reveal this fact, but it has been repeatedly demonstrated by carefully conducted investigations. In addition to the specific automatisms which are required for the four fundamental operations with integers, a number of other automatisms are required for the operations with fractions both common and decimal. At present we have only partial analysis of the examples in these fields, and for that reason it is not possible to state what types of examples are within the range of school work.

The significant characteristics of these abilities or automatic responses are the rate or speed of performance and the accuracy of the response. Thus, the measurement of arithmetical abilities involves determining only at what rate a pupil is able to do examples of the elemental types, and how accurate his answers are. This is accomplished by having him do examples of a given type for a specified time.

From his test paper his rate and per cent of examples correct may be determined. These two quantities represent the measure of his ability to do this type of example.[1]

Limitations of Series B. A complete and detailed measurement would require that a test be provided for each type of example, but fortunately certain combinations can be made. An example in addition consisting of three columns of nine figures each includes the addition combinations, simple column addition, and carrying. Thus, if a pupil responds satisfactorily to examples of this type, we know that he possesses the ability to do the types of addition examples involved therein. On the other hand, if his response to this type of example is unsatisfactory, we do not know just what elemental ability he lacks. The use of a single test of this type, such as Series B, to measure a group of arithmetical abilities has this very obvious limitation in diagnosing the conditions which exist, but it does provide a very satisfactory general survey.

Diagnostic Tests. Any series of classroom tests must not require a large amount of time if they are to be used by any besides the most enthusiastic workers. Thus, in constructing a series of diagnostic tests in the operations of arithmetic, due regard must be had for the amount of time that will be required. Bearing this fact in mind the writer has devised a series of twenty-one tests which require only thirty-one minutes of working time and which it is believed furnish a reasonably complete diagnosis of the abilities of pupils to do the operations of arithmetic with the exception of the types of examples involving mixed numbers and integers with fractions. In order to avoid an excessive number of tests it has been necessary to include more than one type of example within a single test. When this has been done the different types are closely related and they always occur in rotation.

---

[1] Strictly speaking, the number of examples done and the per cent of examples correct is a measure of the pupil's performance rather than of his ability. A pupil's performance is affected by many factors such as his emotional status, physical condition, light, temperature, and the like. Or, it may be that a pupil does not try to do his best on a given test. A pupil's ability can only be inferred from his performance, but when conditions are properly controlled, such inference is reliable in all except a few cases. In order to avoid an awkward form of statement and because the practice is general, we shall speak of a score as a measure of a pupil's ability.

The following samples from each test will illustrate the types of examples included in the several tests:

Addition

Test I

$$\begin{array}{r}4\\7\\2\\\hline\end{array}$$

Test V

$$\begin{array}{r}7862\\5013\\1761\\5872\\3739\\\hline\end{array}$$

Test VII

$$\begin{array}{r}7\\6\\6\\5\\0\\5\\1\\8\\7\\3\\13\\1\\2\\\hline\end{array}$$

Test XII

1/6 + 1/3 =
5/6 + 1/2 =
3/10 + 3/5 =
5/9 + 2/3 =

Test XV

1/6 + 3/5 =
3/12 + 5/8 =

Subtraction

Test II

$$\begin{array}{r}37\\5\\\hline\end{array}\qquad\begin{array}{r}94\\8\\\hline\end{array}$$

Test IX

$$\begin{array}{r}739\\367\\\hline\end{array}\qquad\begin{array}{r}1853\\948\\\hline\end{array}$$

Test XIII

3/4 − 2/5
5/6 − 3/4

Multiplication

Test III

$$\begin{array}{r}6572\\6\\\hline\end{array}$$

Test VIII

$$\begin{array}{r}4857\\36\\\hline\end{array}\qquad\begin{array}{r}560\\37\\\hline\end{array}$$

Test X

$$\begin{array}{r}807\\59\\\hline\end{array}\qquad\begin{array}{r}617\\508\\\hline\end{array}\qquad\begin{array}{r}840\\80\\\hline\end{array}$$

Test XIV

2/3 × 3/4
2/5 × 3/7
5/12 × 3/5

Test XVIII

$$\begin{array}{r}657.2\\.7\\\hline 46004\end{array}\qquad\begin{array}{r}67.50\\.03\\\hline 20250\end{array}$$

Test XX

$$\begin{array}{r}487.5\\.62\\\hline 302250\end{array}\qquad\begin{array}{r}57.28\\9.5\\\hline 544160\end{array}$$

In Tests XVIII and XX the pupil is simply to insert the decimal point in the product which is given. In the samples only the variations in the multiplier are given. Each multiplier is used with three types of multiplicands (657.2, 65.72, 6.572). Thus each test includes six types of examples.

Division

Test IV	*Test VI*	*Test XI*	*Test XVI*
$8\overline{)3840}$	$82\overline{)3854}$	$47\overline{)27589}$	$2/5 \div 1/3$ $4/7 \div 2/3$ $3/8 \div 2/3$

Test XIX		*Test XVII*		*Test XXI*	
$.4\overline{)148}$	*Ans.* :37	$.03\overline{)16.2}$	*Ans.* :54	$.47\overline{)2758.9}$	*Ans.* :587
$.9\overline{)65.7}$	*Ans.* :73	$.07\overline{)1.82}$	*Ans.* :26	$8.2\overline{)38.54}$	*Ans.* :47
$.6\overline{)1.68}$	*Ans.* :28	$.05\overline{).415}$	*Ans.* :83	$79\overline{)36.893}$	*Ans.* :467
$.7\overline{).301}$	*Ans.* :43	$.06\overline{)7.44}$	*Ans.* :		

Test XI is a composite test involving the four "cases" of long division given by Courtis. In Tests XVIII, XIX, and XXI the pupil is to write the answer in the proper place and insert the decimal point. In Test XXI each of the three types of divisor is placed with each of four types of dividends thus providing twelve types of examples.

Marking the test papers and tabulating scores. The test papers are marked in the same way as Series B. The scores are also recorded in the same way. Detailed directions accompany the tests.

Standards. Only tentative standards are available for these tests and they are likely to be changed somewhat in the future. For that reason they are not reproduced here. The best available standards are always furnished with the tests.

Graphical representation. A plan of graphical representation which makes very easy the interpretation of the scores of this series of tests will be given in the next chapter in our consideration of diagnosis. See Figs. 25 and 26.

Summary. In this chapter we have described two series of tests for measuring the abilities of pupils to perform the operations of arithmetic; Courtis's Standard Research Tests, Series B, and Monroe's Diagnostic Tests. The former is to be used for general measurement, the latter for diagnostic measurement. In describing these tests we have shown that ability to do the operations is specific. In the next chapter the meaning of the scores and correctives for the defects discovered by using the tests will be considered. In Chapter VI tests for measuring ability to solve problems will be described.

QUESTIONS AND TOPICS FOR STUDY

1. What is a general test? A diagnostic test?
2. When would you use a general test? A diagnostic test?
3. What is diagnosis?
4. How is the median calculated?
5. What is the "efficiency" of a class?
6. What are the "dimensions" of abilities to do examples?
7. What do we mean by saying abilities to do examples are specific?

CHAPTER V

DIAGNOSIS AND CORRECTIVE INSTRUCTION IN ARITHMETIC

Purpose of giving standardized tests is to furnish basis for improving instruction. As in the case of reading, the fundamental purpose of giving standardized tests in arithmetic is to secure information which the teacher may use in improving the instruction. Series B can be used to secure general information concerning the abilities of the members of a class. The scores will tell the teacher, for each of the four operations, whether his pupils are above or below standard in rate of work and in accuracy. With this information at hand, the teacher knows where he should place the emphasis in his instruction; that is, whether he is devoting insufficient time to practice upon the operations, or whether the pupils are being drilled when they are already up to standard, or whether he should place more emphasis upon the rate of work or upon accuracy.

However, when we recall the nature of ability in the operations of arithmetic, it appears that Series B cannot furnish as complete information as is necessary for planning details of instruction, such as the types of examples which should receive emphasis. A more elaborate series, such as Monroe's Diagnostic Tests, is necessary. In certain cases it is very helpful to supplement the diagnostic tests with an analytical diagnosis.

In this chapter we shall consider (1) the meaning of scores obtained by using Series B and the instruction which should be given to correct the conditions revealed; (2) the use of the Diagnostic Tests and how to use the results; (3) supplementary or analytical diagnosis.

I. COURTIS'S STANDARD RESEARCH TESTS, SERIES B

Type I. Below standard in both rate and accuracy. The causes. There is given in Fig. 21 the record of a fifth-grade class for the addition test. The standards have been indicated by drawing heavy lines. The four divisions into which these lines divide the record sheet make the interpretation of the groups more simple. The records of this class for the other three tests are similar, showing that it is rather conspicuously below standard in both rate and accuracy. Several causes for this condition are possible. The pupils may not know the tables. They may not have learned a good method of work. They may lack sufficient effective drill upon these types of examples. In the absence of scientific information on this point it is the writer's opinion that the last-mentioned cause is the most probable. Saying that the most probable cause is insufficient drill does not necessarily mean that not enough time has been given to drill. Much time may be given to drill and the drill not be effective because the teacher uses a poor classroom procedure.

An illustration of inefficient drill. Frequently the writer has visited classes in arithmetic which were being drilled upon the fundamental operations. A fairly uniform procedure was followed. The same example was dictated to all of the pupils, regardless of whether they needed drill upon this particular type of example or not. Naturally some pupils finished very quickly, and, as they waited for their classmates to finish, there was a tendency for them to become disorderly — a perfectly natural tendency. When a majority of the class had finished the example, the teacher stopped the work and read the correct answer. The process was then repeated. The result was that those pupils who worked slowly completed few, if any, examples during the entire period, and, therefore, received little satisfactory

Test No. 1. Addition.

Score	Score in number of Examples Attempted																									Total	%
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
% 100	—	1	2	3	4	5	6	7	8	9 1	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	
90	Figures in small type indicate scores in number of examples right.		—	—	—	—	—	—	—	—	9	10	11	12	13	14	15	16	17	18	18–19	19–20	20–21	21–22	22–23		
80			—	—	—	4	5 2	6	7 1	8	8	9	10	11	12	12–13	13–14	14–15	15–16	16–17	16–17	17–18	18–19	19–20	20–21	3	
Accuracy 70			—	—	3 1	—	—	5 1	6 1	7 1	7	8	9	10	10–11	11	12	12–13	13–14	14–15	14–15	15–16	16–17	17–18	17–18	4	
60					—	3	4 1	—	5 1	6	6 2	7 1	8	8–9	8–9	9–10	10–11	11	11–12	12–13	12–13	13–14	14–15	14–15	15–16	5	
50			1	—	2	—	3	4	4 1	5 1	5	6	6–7	7	7	8	8–9	9–10	9–10	10–11	10–11	11–12	11–12	12–13	12–14	2	
0 to 49	0	0	0 1	0–1 1	0–1 2	0–2 5	0–2 4	0–3 2	0–3	0–4 1	0–4	0–5	0–5	0–6	0–6	0–7	0–7	0–8	0–8	0–9	0–9	0–10	0–10	0–11	0–11	16	
Total			1	1	3	5	7	3	4	4	2	1														31	

Median Scores: *Rate* 6.9 *Accuracy* 50 *Efficiency* 3 *Standard Scores* ______

FIG. 21. (*Type I.*) THE RECORD SHEET OF A FIFTH-GRADE CLASS IN ADDITION. BELOW STANDARD IN BOTH RATE AND ACCURACY.

drill. The bright pupils spent a considerable portion of their time waiting on the other members of the class, and probably did not need the particular kind of drill which they received.

Some teachers spend a great deal of time in having examples explained by pupils, the explanation consisting merely of an oral reproduction of the process of adding, subtracting, etc. There is a time in the learning process when pupils need explanation, but in the operations of arithmetic, after they understand what to do, attentive repetition is required. This requires an efficient plan of drill. Another point which should be noted is that the drill must be upon all the types of examples which they are to learn to do and not upon the tables or some other single type.

The remedy. (*1*) *A modified classroom procedure.* The type of class instruction described on page 120 can easily be modified so as to make the drill more efficient and to insure that the slow-working pupils will get some satisfactory drill. Instead of dictating only one example at a time, the teacher can dictate several, and stop the work as soon as a few of the faster workers have finished. The slow-working pupils will have some examples completed and the faster workers will not have been idle.

(*2*) *Rate of work must be recognized.* The teacher must recognize that the rate at which the pupil performs the operations is important, as well as the accuracy. This means that, in teaching, the teacher must obtain a measure of the pupil's rate, as well as a measure of his accuracy. If examples are dictated in groups, and the work stopped as suggested in the above paragraph, the number of examples which the pupil does during the class period is a measure of his rate of working. The per cent correct is a measure of his accuracy.

(*3*) *Motivation.* Another means of increasing the scores of a class is to secure a strong motive for the work. Arithmetic is one of the best liked of the drill subjects. This is particularly true of the operations. This being the case, the motivation of drill in arithmetic generally is a comparatively simple matter, and in most cases it will be sufficient simply to start the pupils to work and to keep the work from lagging. When more than this is necessary, the teacher must demonstrate her resourcefulness by providing an effective method or device for the motivation of arithmetical drill. In the lower grades the playing of certain games provides practice upon certain types of examples. In the upper grades ciphering-matches, or, better, the setting of definite standards in both rate and accuracy, are very effective motives.

Standards used to motivate work in arithmetic. The writer has visited classrooms in which the teacher had posted a chart giving the median scores of the class at the beginning of the school year and the standards which should be reached by the end of the year. Teachers testify that this is an effective means of stimulating interest. Figures in the latter part of this chapter illustrate plans for representing the standing of a class on a chart.

Type II. Below standard in rate with satisfactory accuracy. Sometimes a class will be found with a satisfactory median score in accuracy, but conspicuously below standard in rate of work. This condition may be due to the fact that, through the neglect of the rate of work by the teacher, the pupils have not been trained to work rapidly. It may be due to the pupils having the habit of applying some check to their work or doing it a second time. Some pupils work slowly because they are not concentrating their attention upon what they are doing. In giving tests the writer has observed pupils stop and look around the room or out of the window, showing thereby that they were working at a very low pressure. Other pupils have been found to work slowly because they have acquired inefficient methods of work. One such method which is frequently found is the use of an elaborate phraseology or formula in performing the operations. (This cause is treated more fully on page 149.)

The remedy for this type of situation is primarily to place more emphasis upon the rate of work in classroom drills. The pupils should be given timed drills and judged upon the basis of their rate of work as well as the accuracy of their results. The modified classroom procedure suggested above makes this possible. If pupils have acquired habits of work that are undesirable, such as counting on the fingers or tapping out sums or repeating the numbers to be added, subtracted, etc., they should be corrected. Also a strong motive will tend to increase the rate of work. In increasing the rate of work, care must be exercised to make certain that the pupils do not become inaccurate. However, as in the case of silent reading, high rate and high accuracy are not incompatible. In fact they frequently go together, and frequently an increase in the rate of work in arithmetic is accompanied by an increase in accuracy.

Type III. Below standard in accuracy with satisfactory or high rate. If one takes the position that the only satisfactory standard for accuracy is one hundred per cent, this case occurs very frequently. If one accepts the general medians of Table XIV as satisfactory standards, it occurs much less frequently. In this discussion we shall accept the general medians as satisfactory standards. A median score in accuracy may be below standard because the test was given to the pupils in such a way that they became excited and worked very rapidly when they were accustomed to work slowly. When it is suspected that this is the cause, the test should be repeated, using a different form, and exercising care not to excite the pupils. Pupils should become accustomed to working under timed conditions by being timed in doing the regular class exercises.

Another reason for a low median score in accuracy may be the lack of sufficient practice upon the types of examples used in these tests. If this is the cause, the remedy is obvious, and the suggestions given above in respect to effective methods of drill will apply.

Type IV. Scores too widely scattered. An illustration is given of this type of situation in Fig. 22. The median rate of this class is above standard and the median accuracy is only slightly below. Thus, as judged by its median scores the standing of this class is satisfactory, but the scores are not closely grouped about the central tendencies. Although the class is relatively small the rates of work vary from four examples to seventeen examples and the accuracy scores are recorded in each interval of the record sheet. This is an illustration of a condition which prevails to some extent in practically every class. Unless the class is very small, there will be a distribution of scores extending over several intervals of the record sheet.

The overlapping of the several grades. When the distributions of the scores for successive grades are compared, a great overlapping is found. Some pupils in the fourth grade make higher scores than a number of the eighth-grade pupils. Table XV shows the distribution of pupils in a certain city according to the number of examples attempted in the subtraction test of the Courtis Standard Research Tests, Series B. An examination of the table on page 126 reveals these facts:

In the fourth grade twenty-three per cent of the pupils reach or exceed the fifth-grade median.

In the fifth grade twenty-three per cent of the pupils reach or exceed the sixth-grade median.

In the sixth grade twenty-four per cent of the pupils reach or exceed the seventh-grade median.

Test No. 3. Multiplication.

	Score	Score in number of Examples Attempted																								Total	%
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
	% 100	—	1	2	3	4	5	6	7 1	8	9	10	11 1	12	13	14	15	16	17	18	19	20	21	22	23	24	2	
Accuracy	90	Figures in small type indicate scores in number of examples right.		—	—	—	—	—	—	—	—	9	11 2	11 1	12	13 1	14	15	16	17	18	18–19	19–20	20–21	21–22	22–23	4	
	80			—	—	—	4	5	6	7	8	8	9	10 2	11 1	13	12–13	13–14	14–15	15–16	16–17	16–17	17–18	18–19	19–20	20–21	3	
	70			—	—	3 1	—	—	5	6	7	7	8	9	10 1	10–11	11 1	12	12–13 1	13–14	14–15	14–15	15–16	16–17	17–18	17–19	4	
	60			—	2	—	3	4	—	5	6 1	6	7 1	8.	8–9 1	8-9 1	9–10 1	10–11	11	11–12	12–13	12–13	13–14	14–15	14–16	15–16	5	
	50			1	—	2	—	3 1	4	4	5 1	5	6	6–7	7	7	8	8–9	9–10	9–10	10–11	10–11	11–12	11–13	12–13	12–14	2	
	0 to 49	0	0	0	0–1	0–1	0–2	0–2	0–3 1	0–3	0–4	0–4	0–5	0–5	0–6	0–6	0–7	0–7	0–8	0–8	0–9	0–9	0–10	0–10	0–11	0–11	1	
	Total					1		1	2		2		4	3	3	2	2		1								21	

Median Scores: *Rate* 12.3 *Accuracy* 78 *Efficiency* 5 *Standard Scores* ______

FIG. 22. (*Type IV.*) THE RECORD SHEET OF AN EIGHTH-GRADE CLASS IN MULTIPLICATION. SCORES TOO WIDELY SCATTERED.

TABLE XV. SHOWING THE DISTRIBUTION OF THE PUPILS OF A CITY ACCORDING TO THE NUMBER OF EXAMPLES ATTEMPTED, COURTIS'S STANDARD RESEARCH TESTS, SERIES B

Subtraction

Grade	*Number of examples attempted*																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*Median*
IV..	1	7	11	17	19	18	10	8	4	4	1	..	..	..	..	..	..	..	..	..	5.7
V..	..	1	3	9	10	18	23	21	8	2	3	2	..	..	..	..	..	..	..	..	7.4
VI..	..	..	1	2	8	9	20	16	14	14	9	1	4	2	..	..	..	..	..	..	8.6
VII..	..	..	1	..	2	6	9	13	12	16	14	12	6	5	2	1	..	..	1	..	10.4
VIII..	..	..	..	1	..	..	9	13	8	19	12	15	8	5	2	3	2	2	..	1	11.0

In the seventh grade forty per cent of the pupils reach or exceed the eighth-grade median.

This condition is merely typical. Three causes may be given: (1) imperfect classification of pupils; (2) native individual differences in the pupils; and (3) the training which they have received. As a rule only a few cases can be accounted for on the basis of imperfect classification when we remember that except where the school has been departmentalized, the pupils have been grouped not only for instruction in the operations of arithmetic, but also for instruction in other subjects, and that pupils who are low in arithmetic may stand high in some other subjects. Native differences remind us that pupils are different and that complete uniformity cannot be secured. It is, therefore, the third cause which concerns us most.

The effect of class instruction not suited to all pupils. The effect of class instruction is shown when a test is repeated after an interval of a few months. In Table XVI the midyear[1] distributions for a certain city on the addition test of

---

[1] The first test was given just after the midyear promotions. The tests were given to about one hundred and fifty pupils in each grade.

TABLE XVI. THE RANGE OF NUMBER OF EXAMPLES ATTEMPTED

(Upper numbers for each grade are for the mid-year test. Lower numbers refer to the May test.)

Addition

Grade	*Total range*	*Range in number of examples*	*Range of middle 50 per cent*	*Range in number of examples*
IV	1–10	10	3.6–5.9	2.3
	1–17	17	5.0–8.3	3.3
V	2–14	13	4.7–7.0	2.3
	3–16	14	6.8–9.0	2.2
VI	2–12	11	5.7–8.4	2.7
	4–24	21	7.2–11.0	3.8
VII	3–24	22	6.9–10.0	3.1
	4–24	21	8.5–12.3	3.8
VIII	4–18	15	8.1–10.8	2.7
	5–17	13	8.9–12.1	3.2

Series B are compared with those for May. With only two exceptions both the total range of the scores and the range of the middle fifty per cent were increased. This fact shows that the instruction in addition which these pupils received was more appropriate for the brighter pupils than for those who were below standard ability. Some pupils acquired abilities far in excess of the standards for their grade, while others remained conspicuously below the standard. This is merely what we should expect, because those pupils who have profited most under the system of instruction may be expected to continue to profit most. Obviously, if our standards are wisely determined, the pupils who are below standard in ability need instruction and those who are conspicuously above standard may spend their time more wisely upon other subject-matter. If this is not feasible, the methods and devices of instruction should be those most appropriate to those pupils who are below standard. If the methods of instruction are unchanged, it is obvious that the pupils who have learned most readily will continue to do so.

The bright pupil should receive consideration as well as the backward pupil. The usual class instruction does not give the bright pupil efficient training. He is not forced to exert himself. Much of the time he is forced to be inactive. Furthermore, in the case of the tool subjects (the operations of arithmetic, reading, spelling, handwriting, and language), training beyond a certain point is not very profitable. In arithmetic the bright pupil should be given problem work rather than additional training upon the operations.

The remedy. Where reclassification of the pupils is possible and appears wise in the light of the achievements of the pupils in other parts of arithmetic and in their other subjects, a closer grouping can be secured by promoting those who are conspicuously above standard. Some may be put back a grade. But it will be found that many pupils will be below standard only in certain items. To handle these cases, individual, or at least group, instruction is required, because it is obvious that all members of the class do not need the same instruction. Before individual instruction can be wisely planned, additional information concerning the members of the class is needed. This may be secured (1) by using diagnostic tests and (2) by analytical diagnosis. We shall therefore treat Type IV more fully under the head of Diagnostic Tests.

Type V. Irregular development. Fig. 23 shows the median scores of two classes in a certain city. In the number of examples attempted, Class A is nearly three examples higher in subtraction and multiplication than in addition. Class B has very high scores in addition, but lower ones in the other operations, particularly multiplication and division. The relative position of the line showing the median number of examples right shows Class A to be inaccurate in all operations except division, where perfect or one hundred per cent accuracy is indicated. On the other hand, Class B, which in general is relatively very accurate in the other operations does not attain perfect accuracy in division. Such irregular or uneven development is not unusual. It is more frequent and more pronounced in individual pupils than in classes.

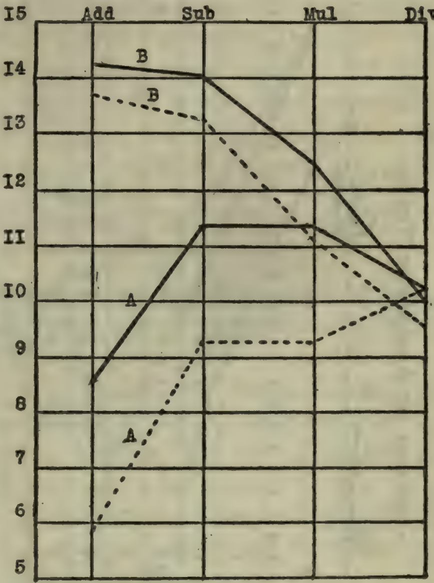
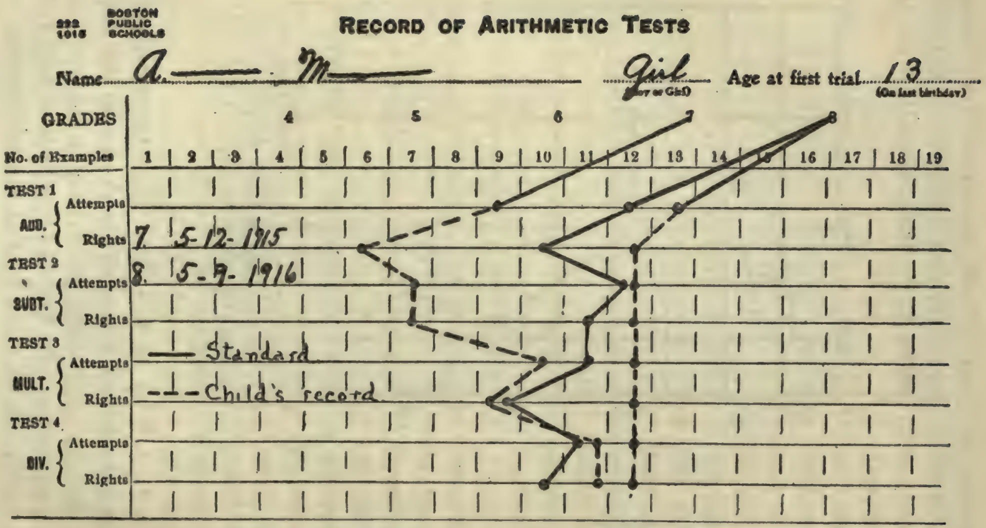


Fig. 23. Showing Median Scores of Two Classes in the Same City. (*After Ashbaugh.*)

In Fig. 24 we show the records of a girl who was tested twice with Series B, first, as she was finishing the seventh grade and again as she was completing the eighth grade. In the figure her records are indicated by broken lines and the position of the eighth-grade standards (Boston) are shown by the solid line. In the seventh grade this girl's development was irregular. In subtraction she had fourth-grade ability (see Table XIV), and in addition she was about fifth-grade. In division she was above eighth-grade ability.

Such uneven development in either a class or an individual pupil is probably caused by not properly distributing the



222
1915
BOSTON
PUBLIC
SCHOOLS

RECORD OF ARITHMETIC TESTS

Name A—— M—— Girl (Boy or Girl) Age at first trial 13 (On last birthday)

7 5-12-1915
8 5-7-1916

— Standard
-+- Child's record

FIG. 24. TWO RECORDS OF ONE GIRL. (*From Ballou.*)

Instructions: On each horizontal line after the word "Attempts," place a dot under the figure representing the number of examples you have tried, and on the line after the word "Rights," place a dot under the figure representing the number of examples that you had right. Beginning with the figure at the top that represents your grade, draw a line through all the dots that you have made.

drill. A study of the median scores of a class on the several tests of the series will tell the teacher where the most emphasis should be placed in the class instruction. A study of the records of a pupil will furnish information for planning individual instruction. In Fig. 24 the broken line drawn for the eighth-grade scores shows that the development of this pupil was evened up during that grade. Such complete evening-up of scores is unusual, but if the teacher places emphasis upon the operations in which low scores are made marked improvement will result.

II. Monroe's Diagnostic Tests

More detailed information is secured by using diagnostic tests. An illustration of the irregular development of a class. By using the diagnostic tests described on page 114, more detailed information can be secured concerning a class or a pupil. In Fig. 25 the median scores of three sixth-grade classes in one city are represented. The plan of graphical representation is similar to that used in Fig. 7. The two scores of a test are represented on the sides of an elongated rectangle, the number of examples attempted on the upper side and the number right on the lower side. The scale on each line has been chosen so that the standard scores of the twenty-one tests for a given grade fall on a straight vertical line. Thus the sixth-grade standards all lie on the vertical line drawn from VI, seventh-grade standards on the vertical line drawn from VII, etc. The scale on the lines has been omitted in order to prevent the crowding of the figure. The numbers and letters at the left of the figure give the number of the test and the operation.

Series B has been used for several years in the city from which the sixth-grade records shown in Fig. 25 were taken, and we may therefore assume that conditions are as good or possibly better than in the average city. Classes A and

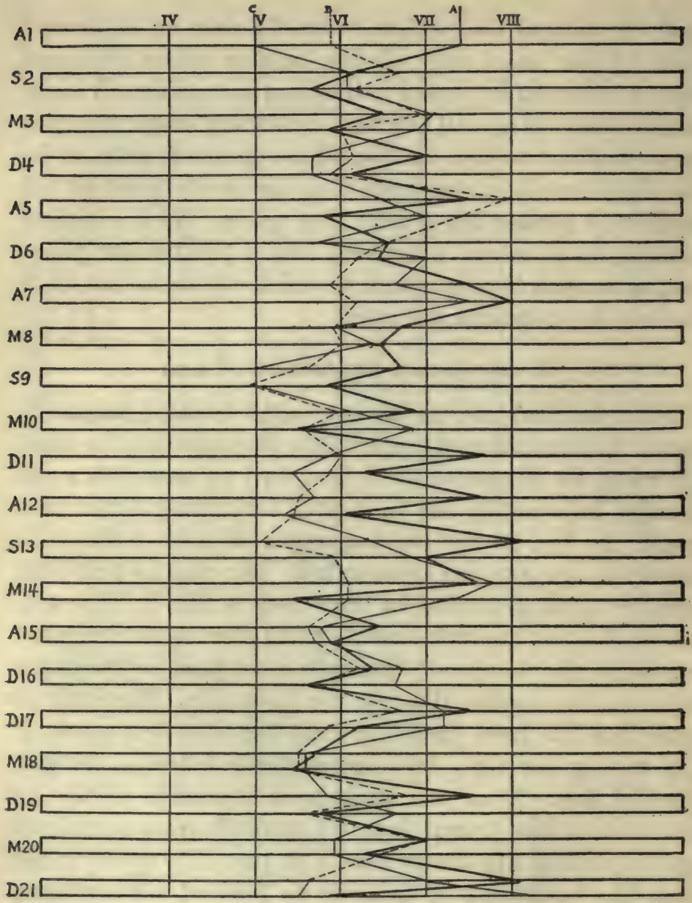


FIG. 25. SHOWING THE MEDIAN SCORES OF THREE SIXTH-GRADE CLASSES.

B have thirty-three pupils each and Class C twenty-three pupils. The figure shows two significant facts: (1) certain points of non-uniformity between the medians of the three classes, and (2) the non-uniform abilities of any one of the classes.

In certain of the tests, such as 2, 6, 8, 15, 18, and 20, the median scores of three classes fall within an interval of about one grade. In others, notably Tests 13 and 21, the extreme difference is much greater. Evidently the teachers of the different classes have not been placing equal emphasis upon the different types of examples. Take, for example, the addition tests (1, 5, and 7). The teacher of Class B has been placing much emphasis upon the type of example in Test 5 (short-column addition with carrying), although it should be noted that she has not neglected the other types of addition examples. On the other hand, the teacher of Class C has neglected short-column addition (Test 1), while the teacher of Class A has given much emphasis to it. In Test 7 there are represented three degrees of emphasis.

The non-uniform character of the abilities of a class is very obvious from the irregularity of the lines representing their abilities. Perfect uniformity would be represented by a straight vertical line. The non-uniformity in the abilities is due to the failure of the teacher to place the appropriate degree of emphasis upon the several types of examples. Some types doubtless require more emphasis than others, and it is the teacher's problem (or is it the problem of the maker of the course of study?) to determine the degree of emphasis which is needed for each type.

An illustration of the irregular development of pupils in the same class. In Fig. 26 there are shown the scores of three sixth-grade pupils selected almost at random from Class A in Fig. 25. H. H. is a twelve-year-old boy, H. C. a ten-year-old girl, and D. H. a girl, age not given. One should expect greater variations when dealing with the scores of individual pupils, but the variations of these scores must be surprising to one who has not studied the subject. Each

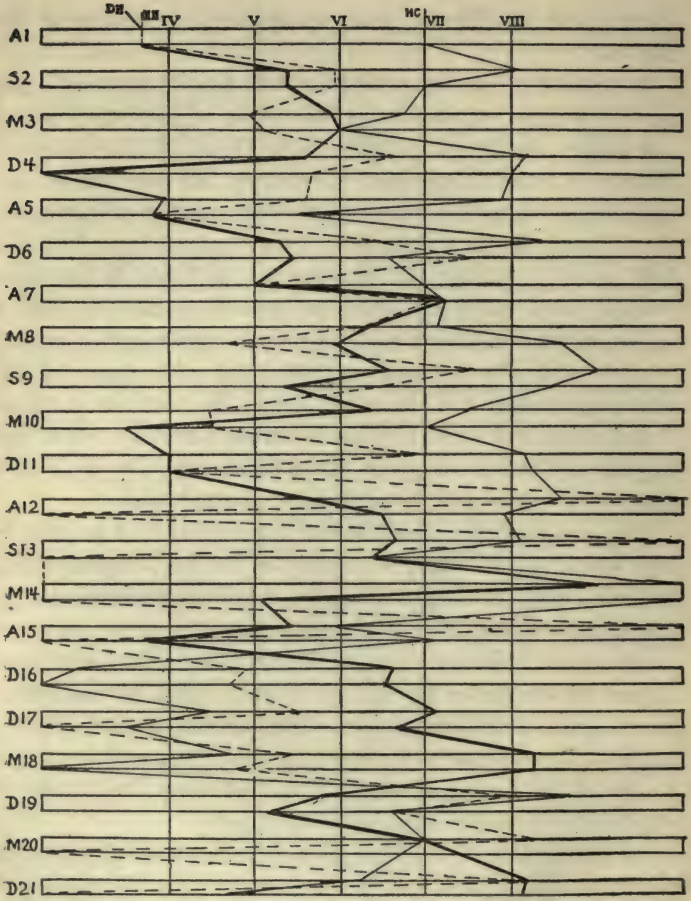


FIG. 26. SHOWING THE INDIVIDUAL SCORES OF THREE SIXTH-GRADE PUPILS IN THE SAME CLASS.

Class A in Fig. 25.

of these pupils has scores on certain tests conspicuously below the standard for the fourth grade and on other tests has scores conspicuously above eighth-grade standards.

Since these pupils were members of the same class, they had probably received the same instruction, and yet the instruction which had been sufficient for one had not always been satisfactory for the others. Pupils differ in the instruction which they need. The instruction which will cause one pupil to learn may make no impression on another.

Arithmetic a complex subject. The facts described above show that the product of arithmetical instruction is complex, much more complex than teachers and supervisors generally realize. The fact that the scores obtained by using these tests show such great variations in the relative degree of ability in the different types of examples, when the pupils have been measured with the Courtis Standard Research Tests, Series B, at regular intervals, is evidence of the need which exists for a series of diagnostic tests. Teachers are probably failing to place the appropriate degrees of emphasis upon the different types of examples, because they are ignorant of what types exist, or do not know the degree of ability which has been attained by their class, and much less the degree of ability attained by the individual pupils. A series of diagnostic tests, such as described in chapter IV, are valuable to the teacher in two ways: (1) as a statement of the important types of examples, and (2) as an instrument for diagnostic measurement.

An illustration of individual instruction. To correct the individual defects individual instruction is needed. A teacher can fit his instruction in the operations of arithmetic to the needs of his pupils by preparing a number of sets of examples, each set being confined to examples of the same type. These sets of examples should be written on cards. Then, instead of dictating examples, the teacher can distribute the cards and have the pupils copy the examples from the cards. If the teacher studies the needs of his pupils, it will be possible for him to distribute the cards so that each pupil will have the type of example upon which he needs practice. The pupil is probably injured by being required to practice upon the wrong type of example, and, hence, it is very important that each pupil be given the type of example upon which he needs practice.

The Courtis Standard Practice Tests used for individual instruction. Courtis has devised a set of Standard Practice Tests,[1] which automatically diagnoses each pupil and furnishes the practice which he needs to remedy his defects. These tests consist of forty-eight sets of exercises, which "have been designed to cover every known difficulty in the development of ability in the four operations with whole numbers." The latest form of these tests (1916) is arranged so that the pupils begin the series by taking Lesson 13, a test involving all types of examples found in the first twelve lessons.[2] All pupils who attain standard ability on this test are excused from the first twelve lessons, because they have demonstrated that they do not need the instruction which these lessons provide. As soon as a pupil who did not attain standard ability on Lesson 13 has finished the first twelve lessons, he takes Lesson 13 again to show that he is now up to standard. Lessons 30, 31, and 44 are also test lessons, and are used in the same way.

Each of the lessons is printed upon a card and a copy is furnished to each pupil. The card is placed beneath a sheet of transparent paper and the example is read through the paper, the work being done on the paper. The lessons have been constructed so that the standard length of time required to complete each one is the same. They are also self-scoring. These two features relieve the teacher of the laborious work of scoring the papers, and make it possible for different pupils to be working upon different lessons at the same time. Thus, when a pupil has demonstrated that he is up to standard on any type of example, he may at once go on to the next lesson. If he is not up to standard on any lesson, his work makes the fact obvious, and he can remain upon that lesson until he acquires the necessary ability without interfering in the least with the work of the other members of the class.

---

[1] Full details regarding these tests may be obtained from the publishers, World Book Company, Yonkers, New York, and Chicago, Illinois.

Another series of exercises, known as the "Studebaker Economy Practice Exercises," and based upon some of the same general principles, has been devised by J. W. Studebaker, Assistant Superintendent of Schools, Des Moines, Iowa. They are published by Scott, Foresman & Company, New York and Chicago. Other series of practice exercises have been devised, but, so far as the writer has examined them, they are less complete and give less promise of efficient means of instruction.

[2] All lessons except the test lessons are confined to a single type of example.

Thus, individual progress is provided for, and at the same time the group formation is retained. A considerable saving of pupils' time is effected by excusing from drill those pupils who demonstrate that they possess standard ability. These pupils can spend this time upon other work.

The use of the Standard Practice Tests in ungraded schools. These Standard Practice Tests also simplify instruction in ungraded schools. In fact they save more time there than in graded schools. The same lessons are used for all pupils in Grades four to eight. Only the time allowed differs. Thus, all of the pupils in a rural school could be instructed at the same time and each pupil receive the practice which he needed.

The most important thing. However, it must not be forgotten that any set of practice exercises is merely a teaching device. It is more important that the teacher explicitly recognize in her thinking that she is instructing a group of pupils who differ widely in native ability, experience, and training, that all do not learn in the same way, and that a limitation should be placed upon training. When she explicitly recognizes these facts, the resourceful teacher will find many devices which will be helpful in adapting the instruction to the needs of the pupils.

III. Analytical Diagnosis

The need for analytical diagnosis. Diagnostic tests, and to a lesser degree general tests such as Series B, locate defects in classes and in individual pupils, but they do not tell the teacher the cause of the defect. The knowledge of the existence of the defect is very helpful to the teacher and she can proceed to eliminate it by increasing the amount of practice or by other devices as has been suggested in the preceding pages. Many cases will be corrected by such treatment, but some will not for the reason that the cause of the defect has not been removed.

The method. Whenever a pupil is found to be conspicuously below standard, the cause should be sought by "analytical diagnosis." This kind of diagnosis includes three types of procedure: (1) observing the pupil as he works; (2) studying his test papers; (3) having him do the examples orally.

Defects discovered by observing the pupil as he works under normal conditions. Courtis[1] recommends this method of diagnosis and lists seven possible symptoms for addition:

1. Child's movements very slow and deliberate, but steady.
2. Child's movements rapid, but variable. Adding accompanied by general restlessness, sighs, frowns, and other symptoms of nervous strain.
3. Child's progress up the column irregular; rapid advance at times with hesitation, or waits, at regular or irregular intervals. Often gives up and commences a column again.
4. Child stops to count on fingers, or by making dots with pencil, or to work out in its head the addition of certain figures.

---

[1] Courtis, S. A., *Teacher's Manual for the Standard Practice Tests* (World Book Company, 1915), pp. 16 ff.

5. Child adds each first column correctly, but misses often on second and third columns.

6. Child's time per example increases steadily or irregularly, particularly after two or three minutes' work; i.e., 15 seconds each for first five examples, 17 seconds each for the next five, 23 seconds for next two, 45 seconds for the next example, etc.

7. Child's habits apparently good and work steady, but answers wrong.

Methods of correcting these defects. Courtis recommends the following correctives for these defects:

1. Slow movements may be due either to bad habits of work or to slow nerve action. In the latter case, the difficulty will prove very hard to control. It is almost certain that no amount of training will ever alter the nerve structure and so remedy the fundamental cause. But in all such cases much can be done to generate ideals of speed, to help the child to eliminate waste motions, and to hold himself up to his best rate.

In any case the procedure would be as follows: Ask the child to add the first example alone so that you may time him. Give him the signal when to start and let him signal when he has finished. Let him make several trials of the same example to make sure that he does not improve under practice. The teacher should then give the child the watch and let him *time the teacher in working the same example.* Comment on difference in child's and teacher's times. Then have the child write in small figures all the partial sums, as shown in the illustration. The teacher should again time the child, letting him *read* to himself the partial sums as rapidly as he can. This will, of course, give the minimum time in which the child could possibly add the example. The time records of a child with true defective motor control will show slight improvement, if any, even with such aid, and probably the only procedure to follow in such cases is to lower the standard to correspond. Where there is a marked difference in time between the original and this last performance, the child will get, for the first time in its life, perhaps, a perfectly clear *conception* of what working at standard speed really means, as well as the *sensation* of really working at that speed. The teacher and child should then practice the same example over and over until the child can *without the crutches* add

$$\begin{array}{r}
{}^{30}46^{15} \\
{}^{26}41^{\ 9} \\
{}^{22}97^{\ 8} \\
{}^{13}60\phantom{^{15}} \\
{}^{7}61\phantom{^{15}} \\
\hline
\end{array}$$

it at the standard rate. Now the teacher can give him the whole test again, urging him to work at his best speed and comparing his results with the first result. The improvement made by ten minutes of this kind of work enables the teacher to say that a proper amount of similar study would produce the changes desired.

But, some teacher will say, "Will the child not learn the example by heart?" This is precisely what is desired. A perfect adder has learned so many examples "by heart" that it is impossible to make up any arrangement of figures that will be in any way new to him. The child in the same way needs to perfect his control over *each* example until he finally attains to mastery over *all*.

2. If the child gives evidence of nervous strain, check his speed, teach him to relax and to work easily and quietly. Get good habits of work first, then bring up speed and accuracy by degrees. The nervousness of a child is usually caused by social conditions, physical health, or temperamental bias. In any event, it is difficult to control. Look out for a large fatigue factor in nervous children.

3. Irregular speed up the column may be due to either of two factors: lack of control of attention, or lack of knowledge of the combinations. The latter factor will be discussed in the following paragraph (4). Attention will be considered here.

There is a limit to the length of time that a person can carry on any mental activity continuously. As time goes on, the mind tends to respond more and more readily to *any* new mental stimulus than it does to the old. The mind "wanders," as it is said. The attention span for many children is six additions, for some only three or four, for others, eight, or ten, and so on. That is, a child whose attention span is limited to six figures may add rapidly, smoothly, and accurately, for the first five figures in the column, giving its attention wholly to the work. As the limit of its attention span is reached, however, it becomes increasingly difficult for it to concentrate its attention. The child suddenly becomes conscious of its own physical fatigue, of the sights and sounds around it. The mind balks at the next addition; it may be a simple combination, as adding 2 to the partial sum, 27, held in mind. It finally becomes imperative that the child momentarily interrupt its adding activity and attend to something else. If this is done for a small fraction of a second, the mind clears and the adding activity will go on smoothly for a second group of six figures, when the inattention must be repeated.

It should be evident that these periods of inattention are critical periods. If the sum to be held in mind is 27, there is great danger that it will be remembered as 17, 37, 26, or some other amount, as the attention returns to the work of adding. The child must, therefore, learn to "bridge" its attention spans successfully. It must learn to recognize the critical period when it occurs, consciously to divert its attention while giving its mind to remembering accurately the sum of the figures already added. This is probably best done by mechanically repeating to one's self mentally, "twenty-seven, twenty-seven, twenty-seven," or whatever the sum may be, during the whole interval of inattention. Little is known about the different methods of bridging the attention spans and it may well be that other methods would prove more effective. The use of the device suggested above, however, is common.

Giving up in the middle of a column and commencing again at the beginning is almost a certain symptom of lack of control of the attention. On the other hand, mere inaccuracy of addition (as 27 plus 2 equals 28) may be due to lack of control over the combinations. If the errors occur at more or less regular points in a column, and if, further, the combinations missed vary slightly when the column is re-added, the difficulty is pretty sure to be one of attention and not one of knowledge.

4. Hesitation in adding the next figure, when not due to attention, is usually due to lack of control of the fundamental combinations. In such cases, however, the hesitation or mistakes are usually repeated *at the same point* on subsequent additions. The teacher should understand that it "takes time to make mistakes," and whenever a lengthening of the time interval occurs, it is a symptom of a difficulty which must be found and remedied.

In this case the remedy is *not* a study of the separate combinations. It has been proved [1] that for most children time spent in study of the tables is waste effort; that the abilities generated are specific and do not transfer. A child may know 6 plus 9 perfectly, and yet not be able to add 9 to 26 in column addition except by counting on its fingers. *The combinations must be learned, of course, but they should be learned by practicing column addition.* Follow the method outlined in paragraph (1) above, having the column added over and over again until both standard speed and absolute accuracy have been attained.

---

[1] See Bulletin no. 2, Department of Coöperative Research, Courtis Standard Tests, 82 Eliot Street, Detroit, Michigan. Price 15c. See also *Journal of Educational Psychology*, September, 1914.

5. The sums of a child who is unable to remember the numbers to be carried, but whose work is otherwise perfect, will usually have the first column added correctly, as well as all single columns. Unfortunately, however, inability to carry correctly is usually a fault of children with weak memories for partial sums in the column. It is well, therefore, to test the carrying habits of any child that is inaccurate. Many children do not add the number carried until the end of the next column; it should, of course, be added to the first figure in the column. If necessary the number to be carried should be emphasized as by saying, when the sum of a column is 27, "carry 2" to one's self as the 7 is written. This is again a time-consuming device which should be adopted only as a last resort. The carrying should be an automatic, unconscious operation. Repeated practice on a few examples until the same become so perfectly familiar that a child's whole attention may be given to establishing correct habits of carrying will prove beneficial.

6. Marked increases in the times required for the successive examples of a test are an indication of a fatigue factor in the control of the attention. Some children are unable to carry on continuously a single activity, as adding, through even a four-minute time interval without a very great loss in power. Two courses are open to the teacher, one or the other of which is sometimes effective: one is to determine the exact length of the interval at which the child can work efficiently, and then try to extend the interval slightly each day; the other is to set the child at work on very long and very hard examples, and to lengthen the time intervals to fifteen or twenty minutes' continuous work. Difficulties of this type are hard to remedy.

Errors discovered by examining test papers. This method of diagnosis cannot always be used because some examples, e.g., the addition test of Series B, are such that the nature of the pupil's errors cannot be determined from his work. However, in common fractions and to a certain extent in subtraction, multiplication, and division of integers, the nature of the errors can be determined.

An illustration of errors made on Series B. Gist[1] examined 812 papers of the Courtis Standard Research Tests, Series B, chosen at random from six schools in Seattle. The frequency, reduced to a per cent basis of each type of error for subtraction, multiplication, and division in the respective grades, is shown in Table XVII. In subtraction "omissions" refer to the number of pairs of digits omitted altogether. Reversions occur when 9 should have been taken from 8, but the digits were reversed. The error indicated by 7–0, is only typical of many similar mistakes when a cipher occurs. The left-hand digit caused some trouble in the eighth grade. In the example:

---

[1] Gist, Arthur S., "Errors in the Fundamentals of Arithmetic"; in *School and Society* (August 11, 1917), vol. 6, p. 175.

$$\begin{array}{r}107795491\\77197029\\\hline 129598462\end{array}$$

the left-hand digit was carried down, as shown.

TABLE XVII. FREQUENCY OF TYPES OF ERRORS IN SUBTRACTION, MULTIPLICATION, AND DIVISION, BASED UPON A STUDY OF 812 TEST PAPERS, COURTIS'S STANDARD RESEARCH TESTS, SERIES B. (GIST.)

	4th	5th	6th	7th	8th
Subtraction:					
Borrowing	54	56	52	51	55
Combinations	36	38	45	44	41
Omissions	2	1	2	3	1
Reversions	1	2	1/2	0	0
7–0, etc.	5	3	1/2	0	0
Left-hand digit	0	0	0	0	2
Multiplication:					
Tables	79	73	73	77	75
Addition	18	20	22	19	20
Cipher in the multiplier	1.5	6	5	4	5
Division:					
Remainder too large	34	39	27	19	10
Multiplication	22	15	19	37	33
Subtraction	11	14	18	25	23
Last remainder 0 and 0 in the dividend	7	15	19	7	11
Multiplicand larger than the dividend	7	4	1	1	1
Failure to bring down all of the dividend	7	4	3	0	6
Failure to bring down correct digit	2	1	4	4	6
Failure to place all of quotient in quotient	7	1	1	3	3
Cipher in quotient as 908–98	3	7	8	4	7

Errors in adding common fractions. The errors which pupils make in the addition of two fractions have been studied so that we know what types are most likely to occur. (1) Counts[1] found in a study of tests given to eighth-grade pupils that sixty per cent of the errors were due to adding the numerators for a new numerator and also adding the denominators for a new denominator, as $3/5 + 1/5 = 4/10$, or $1/9 + 5/9 = 6/18$. It will be noticed that these examples constitute one of the simplest cases in addition of fractions. (2) Twenty-seven per cent of the errors were due to multiplying the numerators for a numerator and multiplying the denominators for a new denominator; as $3/5 + 1/5 = 3/25$, or $1/9 + 5/9 = 5/81$. (3) In a test where it was necessary to reduce the sum to the lowest terms and to a mixed number, Kallom[2] found that nineteen per cent failed to reduce the result to a mixed number and eighteen per cent failed to reduce it to its lowest terms. About half of these pupils failed to make either reduction. About one pupil in twenty failed to express the result correctly when reducing a fraction to its lowest terms, writing $20/15 = 1\ 5/15 = 1/3$, instead of $1\ 5/15 = 1\ 1/3$. (4) Kallom also found certain methods of addition which waste the pupil's time and tend to introduce errors:

> Approximately *one third* found it necessary to reduce the fractions to a common denominator in the first test when the fractions were already similar. Some of these children wrote the fractions over a common denominator, using for a common denominator the denominator of the similar fractions. Others, not noticing that the fractions already had a common denominator, used some multiple,

---

[1] Counts, George S., *Arithmetic Tests and Studies in the Psychology of Arithmetic* (Supplementary Educational Monographs, no. 4, University of Chicago Press), p. 65.

[2] Boston Document no. 3. (1916.) *Arithmetic; Determining the Achievement of Pupils in the Addition of Fractions.* Bulletin no. 7 of the Department of Educational Investigation and Research, p. 19.

making necessary reductions. For example, many children added 3/14 and 1/14 by reducing the fractions to a common denominator of 196. In many cases they then made errors in their work, thus obtaining an incorrect answer to the example. Even if carried through correctly, this is an ineffective and wasteful way of doing such examples.

Another method used by many individuals consisted of finding the least common denominator of such fractions as 1/8 and 3/16 by finding the least common multiple of the denominators by short division as taught in many of the arithmetics. In such cases the following was found:

$$\begin{array}{l}
2\underline{)8 - 16}\\
2\underline{)4 - 8}\\
2\underline{)2 - 4}\\
2\underline{)1 - 2}\\
1 - 1
\end{array}
\qquad\qquad 2 \times 2 \times 2 \times 2 = 16$$

Errors in subtracting common fractions. Counts found two errors which occurred very frequently in subtracting fractions having like denominators. (1) Numerators subtracted for the new numerator and the denominators subtracted for the denominator as $6/9 - 4/9 = 2/0$, or $3/7 - 1/7 = 2/0$. (2) Numerators multiplied for the new numerator and the denominators multiplied for the denominator as $6/9 - 4/9 = 24/81$, or $3/7 - 1/7 = 3/49$. A considerable number of pupils added when subtraction was indicated by a minus sign. This may have been due in part to the fact that both addition and subtraction were included in the same test, but the writer has found similar errors when the two operations were in separate tests.

Correctives for these errors. The first essential for the correction of a defect in a pupil is the knowledge of its existence and nature. Without this knowledge attempts to correct it must be a random trying of methods and devices in hopes that some one will meet the need. Frequently the teacher who knows just what defects exist will be acquainted with some method or device which will serve as an effective corrective. If he is not, a knowledge of the laws of habit formation, which is the type of learning involved in the operations of arithmetic, will help.

The laws of habit formation. Stated in psychological terms, the *first law* is that in the beginning the attention of the learner shall be focalized upon the habit to be acquired. In terms of schoolroom practice this means that the learner shall understand what reaction is to be made to a given stimulus, and shall then react to it in the appropriate manner. This gives the learner the right start.

The second law is that the accomplishment of the step outlined in the first law shall be followed by attentive repetitions. It is not sufficient that there be simply repetitions or drill. The drill must be attentive. In the case of the operations of arithmetic this drill may be detached from the solving of problems, or it may be given in the solving of problems.

The third law states that no exception shall be permitted until the habit is firmly established, which means that the attentive practice must be continued until the operation has become a habit; that is, has been made automatic.

"Borrowing." Table XVI shows that, for the pupils of Seattle, "borrowing" is the most frequent error in subtraction. Some teachers insist that this error can be corrected by using the Austrian or additive method of subtraction instead of the traditional or "take away" method. Although we do not have enough information in order to be able to say positively which method is superior, it appears that the "take away" method is superior to the Austrian method. The latter method may be helpful to certain pupils who have difficulty with subtraction. There are two types of errors in connection with "borrowing." (1) A pupil may fail to "borrow" when he should. (2) "Borrowing" may become a mechanical feature of subtracting and the pupil will "borrow" when the example does not require it. In the first case the pupil must be taught to "borrow." If he has difficulty in grasping the idea, the additive or Austrian method may be presented. It may be that the first law of habit formation has been fulfilled and the pupil needs "attentive repetition" or drill. It will also be helpful to have the pupil do an example several times before proceeding to another. In the second case it will be helpful to give some examples in which no "borrowing" is required. This will demonstrate to the pupil that "borrowing" does not always occur in subtraction.

Combinations or tables. Errors in combinations were next to "borrowing" in frequency, and errors in tables stand at the top of the list in multiplication. Such errors may occur because the pupil does not know certain combinations or because he does not know them well enough. That is, the defect may occur in either the first or second law of habit formation. A few pupils have great difficulty in learning certain combinations. When this is known to be the case, these combinations should be singled out and be made a matter of special drill. Generally when errors in the tables occur in the fourth grade and above, the combinations should not be practiced upon separately, but as they occur in examples. The situation is the same as in addition. (See page 141.)

Remainder too large in division. Outside of errors in multiplication and subtraction this is the most frequent error in division. This error, as well as most of the other errors listed for division, is due to an imperfect plan of procedure; that is, to the failure to apply simple checks at certain steps of the work. The process of division is peculiar in that it is possible to avoid most errors by applying simple checks at certain stages of the work, and pupils should be definitely taught to do this. It is very simple to note whether the remainder is too large by comparing it with the divisor, and this comparison should be taught as a regular step of division.

Failure to reduce sum in addition of fractions. This error is also due to an imperfect routine. Pupils should be taught that the reduction of the answer to a mixed number and to lowest terms is always a part of the work when possible. These errors can be corrected in most cases by the teacher insisting that the reduction be made and providing practice in doing it. Practice upon reduction apart from addition will not be as effective as practice on it as a part of addition. It is a good plan to give no credit for work which is not complete, for the third law of habit formation, permit no exceptions, applies here.

Incorrect methods in handling fractions. Doing such things as adding the denominators when adding fractions is due to the pupil having not fixed the procedure for addition. To correct such defects, the pupil should be shown the correct procedure and drilled upon it. This drill should at first be upon only one type of example, but later lists of mixed types should be used. Here is a good opportunity to use lists written on cards which may be distributed. This makes it convenient to time the pupils, and this is very important because it requires the pupil to decide quickly upon the method to use.

Time-wasting methods. Kallom (p. 144) reports a number of pupils using methods which require more time than is necessary. This will be corrected to a large extent by the teacher emphasizing the rate of work as being important. Of course, it is more important that a pupil have his work correct than to work rapidly, but frequently it is helpful to emphasize the rate of work.

Analytical diagnosis by the oral method. By having a pupil do examples orally, it is possible to discover (1) particular errors in the combinations and (2) imperfect and wasteful methods.

Illustrations of wasteful methods. By using this method of diagnosis the writer has found that in addition many pupils repeat each number to be added. For example they say, "7 and 6 are 13 and 5 are 18 and 4 are 22 and 5 are 27," instead of simply calling the partial sums, as "13, 18, 22, 27." Similar elaborate phraseology is used in the other operations. No error is involved in this method, but it consumes time.

Uhl's oral diagnosis of errors. Using Lesson 1 of the Courtis Standard Practice Tests, which consists of single-column addition, three figures to the column, Uhl[1] studied the errors of pupils by having them do the examples orally and by asking them detailed questions when their method was not made clear. To illustrate the method and also the nature of the defects revealed we quote from his report:

> The findings as to methods employed by pupils in "difficult" combinations is both interesting and significant. The following methods were found in the work of pupils who were tried out in the manner just described. A fourth-grade boy showed by slow work that the combination 9–7–5 was difficult for him. When questioned, he showed that he used a common form of "breaking-up" the larger digits. In working the problem, he said to himself: "$9 + 2 + 2 + 2 + 1 = 16$ and 21." This shows that the 9–7 combination was not known, but that the 16–5 combination was, inasmuch as he arrived at "21" directly after having combined the other two numbers. Another boy of the same grade showed the same type of difficulty in a more pronounced form. He added 8, 6, and 0 as follows: "First take 4, then take 2, then add 8 and 4 makes 12, and 2 makes 14." In adding 9, 7, and 5 he said: "9 and 3 is 12 and 4 is 16 and 2–18; and 2–20; and 1–21." He broke into parts even so easy a problem as $3 + 4 + 9$, adding $9 + 3 + 2 + 2 = 16$.

---

[1] Uhl, W. L., "The Use of Standardized Materials in Arithmetic for Diagnosing Pupils' Methods of Work"; in *Elementary School Journal* (November, 1917), vol. 18, p. 215.

A pupil from the fifth grade presented a quite different method of adding. In adding 4, 9, and 6 she explained: "Take the 6, then add 3 out of the 4. Then 9 and 9 are 18, and 1 are 19." Other problems were worked out similarly: one containing 3, 9, and 8 was solved as follows: "8 and 8 are 16 and 3 are 19 and 1 are 20"; 5, 6, and 9 as follows: "6, 7, 8, 9, and 9 are 18 and 2 are 20." This tendency to build up combinations of 8's or 9's continued in the case of another problem: 6, 5, and 8 were added thus: "6, 7, 8, and 8 are 16 and 3 are 19." Probably her first problem was worked similarly, but I had to have her dictate her method twice before I understood; she then gave it as quoted.

Methods which are quite as clumsy are found in the case of subtraction. One boy of the fifth grade was found to build up his subtrahend in the case of many problems. For example, in subtracting 8 from 37, he increased his subtrahend to 10, then obtained 27, and finally added 2 to 27 to compensate for the addition of 2 to 8. Likewise, in subtracting 7 from 30, he added 3 to 7 and proceeded as before. This boy knew certain combinations very well, but did problems containing other combinations by a method much harder than the correct one.

Even greater resourcefulness was shown by a fifth-grade boy who found the differences between some numbers by first dividing, then noting the remainder or lack of one, then multiplying, and finally adding to or taking from the result as necessary. For example, in subtracting 9 from 44, he proceeded as follows: "Nine goes into 44 five times and 1 less; 4 times 9 are 36, minus 1 equals 35." That is, this boy knew certain multiplication combinations better than he did certain subtraction processes; therefore, he used multiplication, making adjustments either upward or downward as demanded by the problem.

The Remedy. Where a pupil simply uses an elaborate phraseology he should be trained to use a simplified method. The plan of timing the pupil and then having the pupil time the teacher, or comparing his rate with the standard rate, will show him that his method is slow. The use of such methods as described above in the earlier grades may be justified, but teachers should make certain that they are replaced by more efficient ones later. It will help to recognize the rate of work as an important "dimension" of the ability to do examples. When a pupil is found who does not know certain combinations, but is doing the examples by an ingenious method of counting, he must be taught these combinations.

An illustration of corrective instruction. A fifth-grade class [1] was given the Cleveland Survey Tests [2] to determine the types of examples which the pupils could not do with standard ability. This information was supplemented by oral diagnosis in correcting the defects revealed. The regular "class instruction was supplemented at other periods in the day by special help to different pupils in the processes in which they were weak, and they were required to work extra examples in those processes after the help had been given. The drill was limited to the four fundamental operations with integers and fractions. At the end of a month of this kind of instruction, the pupils were given the original test a second time."

In Fig. 27 the record of one pupil, who made some very low scores on the first test, is shown. The figure is so drawn that if the pupil were just up to standard in each test (A, B, C, D, etc.), the line representing his record would be a straight horizontal line. The solid line represents his first record and the broken one his second record. Corrective work was attempted in Tests A, B, F, I, N, and O, and the figure shows a marked improvement was made in these tests. By planning the corrective instruction more carefully it is probable that a still more uniform development might have been obtained. However, as it is we have a striking illustration of what can be accomplished when a diagnosis of a pupil is made and the instruction is based upon the diagnosis.

---

[1] Smith, James H., "Individual Variations in Arithmetic"; in *Elementary School Journal* (November, 1916), vol. 17, p. 198.

[2] These tests are similar to Monroe's Diagnostic Tests except that they contain no tests on decimal fractions.

Summary. In this chapter we have considered the causes of and corrections for the following types of class scores; (1) below standard in both rate and accuracy; (2) below standard in rate with satisfactory accuracy; (3) below standard in accuracy with satisfactory or high rate; (4) scores too widely scattered; (5) irregular development. In connection with the case of irregular development the value of diagnostic tests was pointed out. In dealing with individual pupils who are below standard analytical diagnosis is helpful for discovering the cause of the defect. Several illustrations of this method of diagnosis have been given together with the correctives for handling each case.

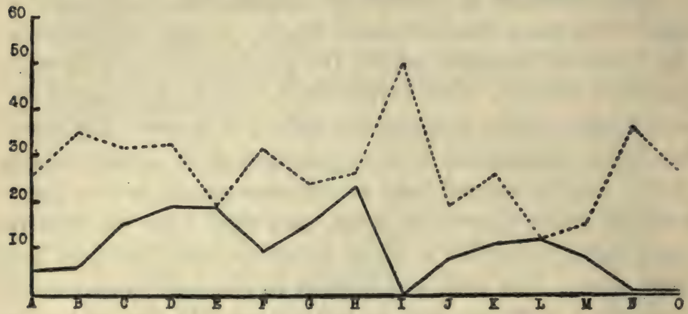


FIG. 27. SHOWING TWO RECORDS OF ONE PUPIL ON THE CLEVELAND SURVEY TESTS.

First record, solid line. Second record, four weeks later; dotted line, corrective work given on Tests A, B, F, I, N, and O.

QUESTIONS AND TOPICS FOR STUDY

1. Do you think pupils will welcome definite objective standards and the use of standardized tests? Why?
2. If you are using standardized tests make charts showing class (or individual) scores in comparison with the standards. Some teachers

have found it helpful to have such charts hung in the classroom. It is also helpful to bring such charts to the attention of the patrons of the school.

3. Make a chart showing how the pupils of your class compare with other classes of the same grade and with classes of other grades.
4. Suppose a pupil is unable to do satisfactorily certain types of examples. How would you proceed to locate his particular difficulties? If you are teaching arithmetic, try out your plan on some of your pupils.
5. What device do you use to provide each pupil with the training which he needs? What devices are suggested in this chapter? Can you suggest additional ones?
6. Pupils who are excused from drill because they do not need it should spend their time doing profitable things. Suggest a number of assignments which might be made to such pupils. The assignments may be in subjects other than arithmetic if it seems wise, but they should be such as not to interfere with the instruction of the other pupils.
7. How do you know that the methods and devices of instruction which you are now using are the best? How could you find out?
8. How do you know that you are not giving too much time to arithmetic? How could you find out?
9. Is a class score which is conspicuously above standard a sign of superior teaching? Why?
10. Construct two tests, each being confined to a single type of example. Give both tests to the same pupils under the same conditions. Compare the two sets of scores.
11. Scientific experimentation will be necessary to determine the best plans of grouping pupils for instruction. These plans are worthy of a trial.
    a. In a building place together for drill those pupils which are most nearly equal in ability as shown by the tests.
    b. Excuse from drill those who have demonstrated that they are above standard.
    c. Have a special "hospital" class for those pupils who have scores conspicuously below standard. A pupil's sentence to the "hospital" would be until he brought his scores up to standard.

CHAPTER VI

THE MEASUREMENT OF ABILITY TO SOLVE PROBLEMS AND CORRECTIVE INSTRUCTION

Description of Monroe's Standardized Reasoning Tests. The measurement of the ability of pupils to solve problems requires a list of problems whose difficulty or value has been determined, because we saw in Chapter I that problems are not equally difficult. Monroe's Standardized Reasoning Tests consist of a series of three tests: Test I for the fourth and fifth grades, Test II for the sixth and seventh grades, and Test III for the eighth grade. Two forms of each test are available so that when it is desired to test the pupils a second or third time, it is not necessary to use the same list of problems. Each problem has been given to several hundred pupils and its value has been determined both for correct principle and for correct answer. Thus, each problem has two values, one a "principle value," which represents the credit to be given for correct reasoning in solving it, and the other a "correct answer value," which represents the credit given for making the calculations correctly when the problem has been worked according to the correct principle.

An important feature of these tests is the manner in which the problems were selected. The writer believes that a satisfactory reasoning test must be composed of problems which are representative in language and content. In order to secure such a list the one- and two-step problems in eight widely used textbooks, which totaled about nine thousand problems, were collected and classified according to the language in which they were stated. The necessity for this classification will be taken up later under the head of "Diagnosis." The types of problems used in the tests occurred in at least five of the eight textbooks, which insures that the language of the problems is representative of that used by the authors of our textbooks.

The tests are printed so that the pupil has space to do each problem on the test paper beside the printed statement of it. This eliminates the necessity of copying either the problem or the work. Thus, the teacher has a complete record of the pupil's work which is valuable for making an analytical diagnosis. The following problems will illustrate the nature of the tests:

A farmer raised 500 bushels of wheat on a field of 40 acres. What was the average yield per acre?	Principle value 2 Answer value 2
A tailor uses $9\frac{3}{4}$ yds. of cloth for a suit. How many yards will it take for 32 suits?	Principle value 1 Answer value 2
A field is 20 rods long, and 12 rods wide. How many rods of fence are needed to enclose it?	Principle value 2 Answer value 1
How much more is earned per day by a man receiving $30 per week than by a man earning $18 per week?	Principle value 3 Answer value 2

Method of giving the tests. Detailed instructions for giving the tests are printed on the test sheets and thus need not be repeated here. One point, perhaps, should be explained. In the case of silent reading and the operations of arithmetic, we have emphasized the fact that ability was "two-dimensional"; that is, that the rate was important as well as the comprehension or accuracy. In solving a problem, the relative importance of rate as a "dimension" of the ability is less, but probably should not be neglected. Accordingly the directions for giving the tests require the pupil to mark the problem which he is working on at the end of a given number of minutes and then he is allowed to finish the test. This gives a rate score and also a score independent of the rate of work.

Scoring the test papers. Instructions for scoring the papers are furnished with the tests, but certain points should be noted here. In order that the "ability to reason" may be measured separately from the "ability to perform operations," each problem is marked for both "correct principle" and "correct answer." A solution is marked correct in principle when it shows that the pupil has reasoned correctly. If a pupil fails to remember correctly some fact, such as the number of pounds in a bushel of wheat or the number of square feet in a square yard, his reasoning is not affected. An answer is counted as correct when (1) the solution is correct in principle and (2) also the answer is numerically correct and in its lowest terms.

Each pupil is given three scores: (1) "Rate of reasoning," which is the sum of the "principle values" of the problems solved within the time limit allowed. (2) "Correct reasoning," which is the sum of the "principle values" of all the problems solved correctly in principle. (3) "Correct answer," which is the sum of the "correct answer values" of those problems which were solved correctly in principle and for which the correct answer was obtained.

Recording the scores. The class record sheet is similar to that used for Monroe's Standardized Silent Reading Tests. A blank for recording the third score is provided. This record sheet and detailed instructions for recording the scores are furnished with the tests. The median scores of the class may be calculated either by the directions given for the Standardized Silent Reading Tests on page 29, or by the directions for Series B on page 104.

Standards. These tests have been used only in a preliminary form, and for this reason standards have not yet been announced. However, standards will be determined as soon as reports have been received from a sufficient number of schools, and any one who is interested may obtain them by writing the Bureau of Coöperative Research, Indiana University, Bloomington, Indiana.

Interpretation of class records. Because the final form of these reasoning tests has not been used, it is not possible to give typical class records upon which to base this discussion of the interpretation of scores. However, the preliminary form was given to over thirteen thousand pupils, and upon the basis of these results types of situations which require correction can be predicted with a high degree of certainty: Type I, low median score for "Rate of reasoning"; Type II, low median score in "correct reasoning"; Type III, low median score in "correct answer" indicating inaccurate calculation; Type IV, scores too widely scattered or distributed.

Type I. Median score for "rate of reasoning" below standard. From the nature of the test it is obvious that the rate of reasoning is not measured separately from the rate of calculation, for a pupil not only "reasons out" a problem, but also performs the necessary operations before he proceeds to the next one. Hence, the "rate of reasoning" score is a measure of the rate of reasoning plus the rate of calculation. For this reason a score may be below standard either because the pupil was slow in his reasoning or because he was slow in performing the operations. This fact must be kept in mind in interpreting this type.

In the preliminary test some classes worked much more slowly than others, apparently because they had not formed the habit of working rapidly. This was probably due to the teacher failing to recognize the rate of work as important. The writer does not believe the rate of work is as important in solving problems as in performing the operations of arithmetic, but it is his judgment that the rate of reasoning should not be neglected. The teacher should at least occasionally time pupils when they are solving problems, telling them, however, that it is more important to have their work right than to solve a large number of problems.

Stone[1] tells of the case of one pupil who had not learned how to work rapidly:

> Pupil, H. C.
>
> Diagnosis: Up to standard in reading ability, did not indulge in undue labeling, physical examination showed no defects, constantly made low scores. Conclusion as to cause of low score: *Mental laziness with lack of realization of the passing of time.*
>
> Treatment: The pupil was first of all made conscious of his status by comparing his score with those of his fellow classmates and with the standard; then he was helped to study his way of working which convinced him of the seat of his difficulty. From day to day lists of approximately equivalent problems were given him with time limit. Much was made of record of scores, gain being expected by both teacher and pupil.
>
> Results: Within a few days notable gain appeared, due to increased ability to direct and hold attention to the work in hand. Contrasted with his previous tendency to wander, the pupil became capable of working continuously in spite of such distractions as people entering the room. After about twenty minutes daily for three weeks he raised his score from 4 to 5.4.[2] Though this is not a large gain in score, the boy had made it largely of his own initiative; he had formed an ideal of concentration, and the concept of *giving attention* to reasoning processes was well under way. It is believed by those who have studied the boy that much of his improvement was due to the convincingness of the objective evidence of his need to improve.

Some pupils worked slowly because apparently they did not know how to think out the plan of solution. They would try one plan and then erase their work or cross it out and try another plan. In such a case the pupils need to be taught how to think. This situation occurs more frequently with individual pupils than with whole classes, and for that reason the corrective measures will be discussed under that head.

---

[1] Stone, C. W., *Standardized Reasoning Tests in Arithmetic and How to Utilize Them.* Teachers College (1916), p. 23.

[2] These scores refer to Stone's Reasoning Tests.

A low median score, due to slowness in performing the operations, may occur in two ways. First, the pupils may not be trained to perform the operations rapidly. This can be verified by giving a test upon the operations such as Series B. If their scores on these tests are below standard in rate, the correctives given on page 123 should be applied. Second, the pupils may be recording their work in some particular form which the teacher requires. The pupils in some classes write out the solution in the form of an analysis or record in other ways which consume time. Orderliness and system are desirable. In a reasonable degree they are necessary, especially when the solution of a problem is long. But it should always be remembered that they are a means or method for making the solution of the problem easier. The teacher should not insist upon a particular form or system when it interferes with the pupil's work.

An illustration of low rate of work due to this last cause and the effect of corrective treatment is given by Stone:[1]

Some pupils of a certain fifth grade

Diagnosis: Many pupils made very low scores, many papers much covered with such statements as, "If one tablet cost 7 cents, 2 tablets . . . etc." Here was evidently one large source of failure.

Treatment: Emphasis was placed on the possibility of saving time by not writing so much, brief labels were devised, originality was encouraged, and approval of pupils and teacher placed on briefest adequate statement.

---

[1] Stone, C. W., *Standardized Reasoning Tests in Arithmetic and How to Utilize Them.* Teachers College (1916), p. 23.

Results: As shown by second test and by daily work, much time was saved for reasoning processes. The following parallel columns show typical results:

Pupil, A. K.

In first test	*In second test*
They would cost $18. If one suit cost $2.50, 9 would cost $2.50 × 9 = $22.50. They would cost $40.50.	$2.50 × 9 = $22.50 $2.00 × 9 = $18. $40.50
Score in first test, 1 1/3.	Score in second test, 3.

Pupil, L. I. C.

In first test	*In second test*
If he sold 4 papers and got twenty cents for them, one half would be 10 cents, and with the other 10 cents he bought Sunday papers, he would buy as many as 2 will go into 10 or 2)10 5 papers.	5 One half would be 10 4 cents and he could buy 20 five.
Score in first test, 3.	Score in second test, 4 1/2.

Type II. Below standard in correct reasoning. In order to understand the reasons for a class being below standard in reasoning and the corrective measures which should be used, it is necessary to understand just what is required of the pupil in solving a problem. The process of solving a problem by reflective thinking may be described in the following steps:

1. It is necessary that the pupil read the statement of the problem with understanding. This is a complex process and involves several abilities: eye-movement, perception, association of meaning with symbols, and combining the several elements of meaning into an understanding of the problem. Out of this should come a definition of the problem, which

is the *first step in reflective thinking*. It should be noted that two kinds of words occur in the statement of problems; *first*, words which describe the setting of the problem or the particular environment in which it occurs, and *second*, words which define quantities or quantitative relationships. This second class of words we may call *technical*. The meanings associated with them must be exact. Take, for example, this problem, "What is the value of sugar obtained at a Vermont sugar camp if it is worth ten cents per pound and six pounds are obtained on an average from each of 1275 maple trees?" Words in this problem such as "Vermont," "sugar," "maple," and "camp" describe the setting. They have nothing to do with the solution of the problem. The technical words are such as "value," "per pound," "are obtained," and "each." They define the relationships which exist between the quantities and are cues for formulating the hypothesis or plan of solution which is another step in the process.

2. Principles applicable to the problem must be recalled. For example, in the problem, "A man invests $893 in some property. He sells the property for $1050. What is his rate of profit?" it is necessary to recall the principle that the rate of profit is calculated upon the amount invested and not upon the selling price. The principles and the meanings of the technical words are the *data or facts which are used in the reflective thinking.*

3. The elements of meaning and the recalled principles are used in formulating a plan of procedure or *hypothesis* concerning the operations to be performed upon the quantities of the problem. In doing this each element of meaning must be given its proper weight. A relatively inconspicuous term may require an operation. For example, in the problem, "A rectangular court 72 feet by 120 feet is to be paved at a cost of $2 per square yard. What will be the ex-

pense?" the use of "square yard" instead of "square foot" in the statement of the problem makes necessary an additional operation.

4. The hypothesis thus formed should be *verified.* Generally this does not occur as an explicit step. It consists of seeing that the hypothesis is in agreement with the several elements of meaning and the recalled principles.

5. The operations outlined in the hypothesis are performed. Strictly speaking, *this is not a step in the reasoning process.* This is completed when a correct plan of action is formulated.

This analysis assumes that the problem is solved by *reflective* thinking. In many cases the pupil does not reflect. If it is very familiar he may automatically identify it as requiring a particular operation or operations. This may happen after only a partial reading of the problem. Under any circumstances this procedure is probably more of the nature of a "short-circuiting" of the reflecting thinking process than an exception to it. When the problem is unfamiliar the pupil may try random guessing at the plan of solution.

It was noted that the data used in solving a problem come from two sources, recalled principles and the meanings of the technical words used in the statement. The ability to associate the correct meaning with one term does not imply the ability to associate the correct meaning with another term. The ability to solve the problem, "At $55 each how much must a farmer pay for 25 cows?" does not make certain the possession of the ability to solve "Find the duty on $600 worth of clocks at 40% *ad valorem,*" although the same operation is required. The technical terms, such as "$55 each," "pay for," "find the duty," and "*ad valorem,*" are sufficiently different so that a pupil might know the meaning of one set without knowing the meaning of the other. The meaning of the technical terms in a problem furnishes important data or cues for the judgment concerning the operations to be performed. In many cases it appears that the determining data come from this source. Thus, in measuring a pupil's ability to solve printed problems we are measuring his knowledge of technical terms as well as his ability to use this knowledge in formulating plans of procedure.

The reading of problems is difficult because many forms of statement are used. The solving of a problem requires a careful reading of it with a high degree of understanding and such reading of problems is a more complex and difficult task than we commonly realize. Problems are stated in many forms and the total "technical" vocabulary which is required of a pupil by the time he completes the work of the elementary school is a large one. For an illustration, take this problem situation: Given $7.50 paid for silk and price per yard $1.50, to find number of yards purchased. Excluding different arrangements of the words used, twenty-eight different forms of statement were found in examining eight textbooks for describing this problem situation, and additional forms could be constructed.

1. How many yards of silk at $1.50 per yard can be bought for $7.50?
2. The silk for a dress cost $7.50. How many yards were purchased at $1.50 per yard?
3. At $1.50 per yard, how many yards of silk does a woman get if the amount of the purchase is $7.50?
4. At the rate of $1.50 per yard my bill for silk was $7.50. How many yards were purchased?
5. How many yards of silk at $1.50 a yard does a bill of $7.50 represent?
6. When silk is $1.50 a yard, a piece of silk costs $7.50. How many yards in the piece?
7. At $1.50 a yard how many yards of silk does a merchant sell if he receives $7.50 for the piece?
8. Mrs. Jones purchased silk at $1.50 a yard. The entire amount paid was $7.50. How many yards were bought?

9. Silk was sold at $1.50 per yard. A check for $7.50 was given in settlement. Find the number of yards bought.
10. At $1.50 per yard, how many yards can be bought for $7.50?
11. A merchant sells a number of yards of silk for $7.50. The price being $1.50 for each yard, how many does he sell?
12. I invested $7.50 in silk at $1.50 per yard. How many yards did I buy?
13. When silk is $1.50 per yard, how many yards can be bought for $7.50?
14. When silk is sold for $1.50 for each yard, what quantity can be bought for $7.50?
15. At the rate of $1.50 per yard, how many yards can be bought for $7.50?
16. Silk is selling for $1.50 per yard, how many yards should be sold for $7.50?
17. At a cost of $1.50 a yard, how many yards can be bought for $7.50?
18. Silk was bought at a cost of $1.50 per yard. At that rate, how many yards can be bought for $7.50?
19. At $1.50 a yard a piece of silk cost $7.50. How many yards in the piece?
20. How many yards of silk at $1.50 can I buy for $7.50?
21. $7.50 was paid for silk at $1.50 per yard. How many yards were bought?
22. Find the number of yards; cost $7.50. Price per yard $1.50.
23. The cost of a piece of cloth is $7.50 and the cost per yard is $1.50. How many yards are there in the piece?
24. A woman paid $7.50 for a piece of silk that cost her $1.50 per yard. How many yards were there in the piece?
25. A woman had $7.50 and bought silk at $1.50 a yard. How many yards did she buy?
26. A quantity of silk at $1.50 per yard cost $7.50. What was the quantity?
27. Silk is $1.50 a yard and I bought $7.50 worth to-day. How many yards did I buy?
28. A woman's bill for silk was $7.50. If each yard cost $1.50, how many yards were bought?

This illustration of the variety of terms which are used in the statement of one problem becomes more significant when we remember that this is just one problem and a relatively simple one. It should be clear that learning to read problems is not an easy matter.

A test to measure a pupil's knowledge of words used in problems. The test given below was devised to measure a pupil's knowledge of the meaning of words used in stating problems. The words in this test were found to be "common" to three or four of the newer textbooks for the grades in which the test was given. A preliminary test was given first to insure that the pupils would understand what the test asked them to do.

VOCABULARY TEST IN ARITHMETIC[1]

Name...................... Grade ....

1. Put *w* beside each word that tells what a man's work is.
2. Put *m* beside each word about money.
3. Put *l* beside each word that might be used about land.
4. Put *i* beside each word that is the name of something to put things in.

basin
salary
retail
mason
basket
lot
bucket
rent
machinist
coins
broker
excavate

area
carpenter
field
bin
real estate
poultry
fares
dealer
tailor
barrel
wages
commission

merchant
cashier
building lot
attend
teamster
jars
debts
gardener
expenses
nickel
owe
schedule

profit
pasture
earn
collect
tank
acre
income
insurance
miller
cistern
customer

In Table XVIII the per cent of pupils in both the fourth and fifth grades who failed to mark the words correctly is given. Thus, forty per cent of fourth-grade pupils and

---

[1] Chase, Sara E., "Waste in Arithmetic," *Teachers College Record* (September, 1917), vol. 18, p. 364.

TABLE XVIII. SHOWING PER CENT OF FAILURES ON VOCABULARY TEST IN ARITHMETIC

	Grade			*Grade*	
	IV	*V*		*IV*	*V*
basin	26%	20%	merchant	45%	24%
salary	40	20	cashier	52	20
retail	30	16	building lot	65	32
mason	78	36	attend	4	4
basket	13	8	teamster	88	40
lot	9	12	jars	40	16
bucket	30	16	debts	91	32
rent	56	28	gardener	17	8
machinist	35	28	expenses	65	20
coins	26	16	nickel	13	16
broker	56	45	owe	91	60
excavate	4	4	schedule	17	4
area	91	36	profit	78	45
carpenter	26	8	pasture	52	32
field	9	8	earn	60	36
bin	43	20	collect	17	8
real estate	100	68	tank	65	36
poultry	40	36	acre	56	32
fares	65	60	income	96	41
dealer	35	28	insurance	91	42
tailor	13	4	miller	26	12
barrel	30	32	cistern	100	88
wages	70	28	customer	35	44
commission	100	68			

twenty per cent of fifth-grade pupils failed to mark "salary" as a "word about money." All of the words in this list are not technical terms, but a number, such as "salary," "rent," "area," "field," and "bin," are used in designating the relationship of quantities in problems. For example, "A man receives $185 per month. What is his yearly salary?" or, "A house rents for $40 per month. How much is that a year?" In the first problem a pupil cannot reason about the situation unless he knows that "salary" refers to the "$185 per month" which the man receives.

In another test pupils were asked to draw the figures named in Table XIX. The numbers in the table are the per cent of pupils in each grade who failed to draw the correct figure.

TABLE XIX. SHOWING PER CENT OF PUPILS WHO FAILED TO DRAW CORRECTLY THE FIGURES NAMED

	Grades		
	III	*IV*	*V*
Square	10	4	8
Rectangle	100	39	80
Triangle	50	35	20
Oblong	15	40	4
Rectangular plot		87	68

These two simple tests show something of what the situation in arithmetic probably is. We are asking pupils to solve problems when they do not know the meaning of the terms used in the problems. We must, therefore, begin to give explicit instruction in the meaning of technical terms.

When a class is below standard in correct reasoning, one of two conditions exists. First is a case of ignorance; the pupils do not know the meaning of the technical terms or cannot recall the required principles and facts. The second may be the lack of knowing how to use this information or the lack of a sufficient motive. After examining the test papers of several thousand pupils, it is the writer's judgment that the first is the more frequent condition, but often pupils fail to reason correctly in solving problems because they have no plan of thinking. We saw in the case of silent reading that one cause of poor comprehension was the failure to verify the meaning obtained. In solving problems pupils "accept" an incorrect solution because they do not verify their plan; that is, they omit the fourth step in the process as outlined on page 162.

The general correctives are suggested by the above analysis. The pupils should be taught the arithmetical meaning of the technical terms used in stating problems. They should also be trained to have a good procedure, to be somewhat systematic in their reasoning. Especially should emphasis be placed upon the step of verification.

Type III. Below standard in calculation. These tests were not designed to measure ability to perform the operations of arithmetic. For this reason too much importance should not be attached to the "correct answer" scores; but when these scores show a class to be conspicuously below standard in calculation, the pupils should be given one of the series of tests described in Chapter IV. If these tests show the class to be below standard, the correctives prescribed in Chapter V should be used. Only one point needs comment here. If a class is found to be up to or above standard when tested on the operations separately, then the teacher has the problem of causing the pupils to use this ability in solving problems. Then the teacher should give less "isolated" drill — that is, drill upon examples — and more practice in the solving of problems.

A frequent source of error is the copying of figures. Some pupils copy the wrong figures, as 85 for 55. Others write all numbers as dollars pointing off two places. Still others, when they wish to subtract 240 from 60000, write

$$\begin{array}{l} 60000 \\ \underline{240} \end{array}$$

Type IV. Scores widely scattered. As in silent reading and the operations of arithmetic, frequently the scores of a class will be found widely scattered or distributed. Some pupils will have relatively high scores, others will have very low scores. The remedy is to give individual or group instruction to those who have low scores. Those who have high scores also need special instruction. It may be that they should devote some of the time which they are now giving to the problems of arithmetic to some other subject. A few cases may be adjusted by a reclassification. The corrective instruction for those below standard can be best presented in connection with certain typical errors.

Neglect of certain technical words. An examination of ninety-five fourth-grade test papers revealed the following solutions of this problem: "How much more is earned per day by a man receiving \$30 per week than by a man receiving \$18 per week?"

Solutions correct in principle	23	pupils
$30 + 18 = 48$	15	"
$30 - 18 = 12$	38	"
$30 \times 18 = 540$	4	"
$30 - 18 = 22$	2	"
$30 - 18 = 28$	2	"
$30 \times 18 = 130$	2	"
\$3.00 and \$5.00	3	"
$30 - 18 = 12/30$ more	2	"
$3018 \div 7 = 4148$	1	"
$30 + 10 = 40$	1	"
$30 \times 18 = 60.6$	1	"

The solutions "$30 - 18 = 12$," "$30 - 18 = 22$," and "$30 - 18 = 28$" indicate that the pupils neglected the technical phrase "per day." If this phrase did not occur in the problem these solutions would be correct in principle. It might be that some of the pupils did not know the significance of this term. Solutions such as "$30 + 18 = 48$" and "$30 \times 18 = 540$" indicate either a complete ignorance of the technical term, "how much more," or failure to reason at all.

In the case of this problem, "A car contains 72,060 lbs. of wheat. How much is it worth at 87 cents a bushel?" many fifth-grade pupils gave no evidence that the number of pounds must be reduced to bushels. In the problem, "What are the average daily earnings of a boy who receives $0.88, $0.25, $1.15, $0.75, $0.50, and $0.60 in one week?" a very large per cent of the pupils failed to pay attention to the word "average." Its presence in the problem requires that the sum of the earnings be divided by 6.

The corrective for the neglect of technical terms is to teach the pupils their meanings. In this case the pupils who simply subtracted 18 from 30 need to be taught that "how much per day" when the amount is given for the week means division by the number of days in the week. When pupils do not know the meaning of "average" they must be taught.

Guessing instead of thinking. An excellent illustration of this type of procedure is given by Adams.[1] It occurred in an English school in what is the equivalent of our seventh grade. The problem, "If 7 and 2 make 10, what will 12 and 6 make?" is not the sort which we are accustomed to, but this fact does not destroy the value of the illustration:

> A look of dismay passed over the seventy-odd faces as this apparently meaningless question was read. Everybody knew that 7 and 2 did n't make 10, so that was nonsense. But even if it had been sense, what was the use of it? For everybody knew that 12 and 6 make 18 — nobody needed the help of 7 and 2 to find that out. Nobody knew exactly how to treat this strange problem.
>
> Fat John Thomson, from the foot of the class, raised his hand, and when asked what he wanted, said: —
>
> "Please, sir, what rule is it?"

---

[1] Adams, John, *Exposition and Illustration in Teaching*, pp. 176–78.

Mr. Leckie smiled as he answered: —

"You must find out for yourself, John; what rule do you think it is, now?"

But John had nothing to say to such foolishness. "What's the use of giving a fellow a count[1] and not telling him the rule?" — that's what John thought. But as it was a heinous sin in Standard VI [seventh grade] to have "nothing on your slate," John proceeded to put down various figures and dots, and then went on to divide and multiply them time about.

He first multiplied 7 by 2 and got 14. Then, dividing by 10, he got 1 2/5. But he did n't like the look of this. He hated fractions. Besides, he knew from bitter experience that whenever he had fractions in his answer he was wrong.

So he multiplied 14 by 10 this time, and got 140, which certainly looked much better, and caused less trouble.

He thought that 12 ought to come out of 140; they both looked nice, easy, good-natured numbers. But when he found that the answer was 11 and 8 over, he knew that he had not yet hit upon the right tack; for remainders are just as fatal in answers as fractions. At least, that was John's experience.

Accordingly, he rubbed out this false move into division, and fell back upon multiplication. When he had multiplied 140 by 12, he found the answer 1680, which seemed to him a fine, big, sensible sort of answer.

Then he began to wonder whether division was going to work this time. As he proceeded to divide by 6, his eyes gleamed with triumph.

"Six into 48, 8 an' nothin' over, — 2 — 8 — 0 an' no remainder. I've got it!"

Here poor John fell back in his seat, folded his arms, and waited patiently till his less fortunate fellows had finished.

James[2] knew from the "if" at the beginning of the question that it must be proportion; and since there were five terms, it must be compound proportion. That was all plain enough, so he started, following his rule:

"If 7 gives 10, what will 2 give? — less."

Then he put down

7 : 2 :: 10 :

---

[1] Scotch: any kind of arithmetical exercise in school work.

[2] The clever boy of the class.

"Then if 12 gives 10, what will 6 give? — again less." So he put down this time

12: 6

Then he went on loyally to follow his rule: multiplied all the second and third terms together, and duly divided by the product of the first two terms. This gave the very unpromising answer 1 3/7. He did not at all see how 12 and 6 could make 1 3/7. But that was n't his lookout. Let the rule see to that.

After examining a large number of test papers, this account appears to the writer to describe the mental processes of a considerable number of pupils. They have not learned to think. The teacher insists that they "try" and they put down figures. They have been taught that it is worse to admit that they cannot solve a problem than to try it by unintelligent guessing. It seems that pupils should be taught to admit frankly that they do not know something when they don't know rather than to try to "bluff."

The corrective. Pupils are frequently taught to solve by rule rather than to reason. Rules are helpful when used properly, but teachers should train pupils to think, to associate definite meanings with technical terms, to combine these meanings and recalled rules into a plan of solution and to verify the proposed solution. To do this, the teacher should at first use simple problems, such as, "What is the area of a field 40 rods by 60 rods?" or, "What is the cost of 15 cows at $60 apiece?" and explain to the pupils that the words, "What is the area," "What is the cost," together with the form of the remainder of the statement of the problem, tell one what operation to perform. The words used in stating a problem when properly understood tell one what the plan of solution should be. Occasionally it is necessary to recall rules or principles, but these are suggested by the words of the problem. If this idea can be impressed upon a pupil, progress will have been made in teaching him to think.

Illustrations of failure to verify answer. Frequently pupils give answers which are absurd, thereby furnishing evidence of failing to apply even a common-sense check to their answers. The following are a few illustrations of this practice:

Problem: "A baker used 3/5 lb. of flour to a loaf of bread. How many loaves could he make from a barrel (196 lbs.) of flour?"

Solution: "3/5 of 196 lbs. = 117 3/5 loaves." This solution was given by a large number of sixth- and seventh-grade pupils. One sixth-grade pupil gave this: "3/5 × 39 = 39 1/3 loaves."

Problem: "At the rate of $4 for an 8-hour day, how much is due a man for 6 1/2 hours work?"

Solution (sixth grade): "13/2 × 4/1 = $26." This solution was given by a considerable number of pupils. Some sixth-grade pupils gave this: "8 − 6 1/2 = $2 1/2." One gave this, "6 × 4 = 24, 24 × 8 = $272 1/2."

The corrective. Some teachers recommend requiring pupils to estimate the answer before beginning the solution. For example, in the first problem above, the pupil could determine whether the number of loaves would be greater or less than the number of pounds of flour, 196. In the second problem, the pupil could determine whether a man would receive more or less for 6 1/2 hours than for 8 hours. This makes a common-sense verification a part of the solution of a problem.

Other reasoning tests. Several other tests have been devised to measure the abilities of pupils to solve problems involving reasoning, but none of them have been widely used. Some years ago Stone[1] worked out a reasoning test which has been used in several cities, and in a number of city school surveys, so that we have rather definite standards as to what may be expected from its use. Starch has devised a test which is called Arithmetical Scale A.[1] This scale included a number of the problems used by Stone, Courtis, and Thorndike. They have been evaluated upon the basis of difficulty and arranged in order of increasing difficulty. The pupils are allowed as much time as they need and a pupil's score is the value of the most difficult problem done correctly.

---

[1] Stone, C. W., *Arithmetical Abilities and Some Factors Determining Them.* Teachers College Contributions to Education, no. 19. (1908.) See also Stone, C. W., *Standardized Reasoning Tests in Arithmetic and How to Utilize Them.* Teachers College Contributions to Education, no. 83. (1916.)

QUESTIONS AND TOPICS FOR STUDY

1. What are the steps in the solving of a problem in arithmetic?
2. To what extent and how is silent reading involved in solving problems?
3. Why must the problems which are used in a test be evaluated?
4. How would you go about teaching a pupil to reason in solving problems?
5. How could you find out whether your pupils are lacking in vocabulary or not?
6. On page 160 why is the form of solution on the second test better than the form on the first?
7. What reasons can you give for the absurd answers and forms of solution which many pupils give to problems? How could you correct these defects?

---

[1] Starch, Daniel, "A Scale for Measuring Ability in Arithmetic"; in *Journal of Educational Psychology*, vol. 7, pp. 213–22.

CHAPTER VII

THE MEASUREMENT OF ABILITY IN SPELLING AND CORRECTIVE INSTRUCTION

Making a spelling test. In order that the method of measuring ability in spelling may be understood, certain things in connection with the making of a spelling test must be explained. The following questions are some which must be considered: (1) What words should be selected for a test? (2) How difficult should the words be? (3) How many words should be used? (4) How should they be given?

(1) Selection of words for a test on the basis of frequency of use. The English language contains many words. Some of these the average person never uses, others he uses only occasionally and a few he uses very frequently. For practical purposes there is no advantage in one being able to spell words which he never uses, and the makers of courses of study and textbooks in spelling are attempting to eliminate these words. Hence, it is obvious that such words should not be used to measure the ability of pupils to spell. Of the other two classes it is more important to be able to spell those words which are used most frequently, and for that reason they should be used in a spelling test if it is most helpful to the teacher. Hence, the first step in the selection of words for a test is to determine what words are used most frequently in written language.

Ayres's determination of the most commonly used words. In determining the most commonly used words, the method employed has been to examine written material of several types, such as letters, newspapers, and children's compositions, and to obtain a list of the words used and the number of times each word occurs. Ayres[1] has combined the results of four such studies. Two of these studies were based on letters, the third upon newspapers, and the fourth upon selections of standard literature. The material examined in the four studies aggregated 368,000 words, written by 2500 different persons.

It was the original intention of Ayres to obtain a list of the two thousand most commonly used words, but this was impossible because the material examined was found to consist of a few words used many times, and of a larger number of words used only a very few times. It was found that fifty different words were used so frequently that they made up approximately half of the material examined. In order to secure a list of the thousand most frequently used words, it was necessary to include words which were found only forty-four times in the 368,000 words of material examined. Other studies have been made to determine the words which are used most frequently in written language, but the resulting lists have not been arranged in a form which is convenient to use for testing purposes. Hence, we shall limit this discussion of the measurement of spelling ability to the one thousand words of Ayres's list which is published with the title *A Measuring Scale for Ability in Spelling*.[2] However, one other study may be mentioned to illustrate further this method of determining the words which should receive attention in teaching spelling and in the measurement of spelling ability.

Jones's list of words used by school-children. Jones collected compositions from pupils in Grades two to eight inclusive. In order that a record of the complete writing vocabulary of each pupil might be obtained, a large number of compositions were written, the number per pupil ranging from 56 to 105. A total of 75,000 themes, consisting of a total of 15,000,000 words and written by 1050 pupils residing in four States, were examined. However, only 4532 different words were used by these pupils. Unfortunately, Jones does not tell us how many times each word was used so that we cannot obtain a list of the words which the children used most frequently.

---

[1] Ayres, L. P., *Measurement of Ability in Spelling*. Bulletin of the Division of Education, Russell Sage Foundation. (New York City, 1915.)

[2] The reader should have a copy of this scale in order to properly understand this chapter. See Appendix for directions for ordering a sample package of tests.

(2) **Determination of the difficulty of words.** After we have a list of the most commonly used words, such as Ayres has given us, there remains the problem of determining the relative difficulty of the several words. It is a well-known fact that some words are more difficult to spell than others.[1] The words included in a test either must be equal in difficulty or their relative difficulties must be known. Otherwise we will be using a measuring instrument consisting of unequal units, but will be considering the units to be equal. Doing this makes our measurements inaccurate. The spelling difficulty of words for a given group of children may be determined by having the words spelled by them. From the per cent of correct spellings of each word the relative difficulty of the words may be calculated. Words which are misspelled an equal per cent of times by pupils of a given grade are equal in difficulty for that group. In the absence of this information it is practically impossible for a teacher to judge the difficulty of the words. Buckingham concluded that the judgment of a single teacher is almost of no value. "It may be good and it may be bad; and it is about as likely to be the one as the other."

---

[1] The spelling difficulty of a word has two interpretations. It may be taken to mean the difficulty which children have in learning to spell it. It may also refer to the frequency with which it is misspelled. The latter meaning will be used in this chapter.

How Ayres determined the difficulty of the words in his list. To determine the words of equal difficulty and the relative difficulty of the groups of words, Ayres divided the thousand words into fifty lists of twenty words each. Each list of words was spelled by the children of two consecutive grades in a number of cities. The thousand words were then divided into another fifty lists of twenty words each. Each of the new lists was spelled by the children in four consecutive grades. In all, 70,000 children spelled twenty words, making a total of 1,400,000 spellings, or an average of fourteen hundred spellings of each of the thousand words.

Upon the basis of this information Ayres classified the words into twenty-six groups, the words of each group being approximately equally difficult for school-children of a given grade.[1] This classified list, together with the per cent of pupils in each of the grades who spelled the words of each list correctly, has been printed with the title, *Measuring Scale for Ability in Spelling*. Strictly speaking, the Ayres *Measuring Scale for Ability in Spelling* is not a measuring instrument in itself, but rather a list of the foundation words of the English language, classified into twenty-six groups according to spelling difficulty. The teacher may use this list as a source of words for constructing spelling tests.

Pupils are not tested when words are too easy. When a pupil spells correctly all of the words of a given list, we do not have a measure of his spelling ability. We simply know that he can spell these words correctly; we do not have any information concerning how far beyond this list his spelling ability extends. In fact, the pupil has been given no opportunity to show how well he can spell. It is a well-known fact that the pupils of any grade or of any class are not equal in ability, but exhibit a wide range of ability.

---

[1] For the details of the method employed see Ayres, L. P., *Measurement of Spelling Ability*, pp. 22–35.

Thus, in testing a class it is necessary to use words for which the average per cent of correct spellings is less than one hundred. Ayres recommends that in making a test for the pupils of a given grade, the words be taken from the column for which an average of eighty-four per cent of correct spellings may be expected.[1]

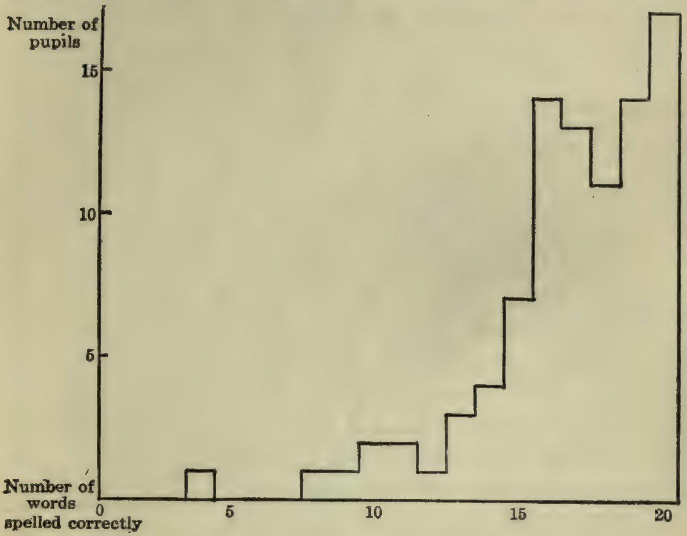


FIG. 28. SHOWING THE DISTRIBUTION OF 91 PUPILS ACCORDING TO THE NUMBER OF WORDS SPELLED CORRECTLY.

Class average, 84 per cent.

Fig. 28 represents a typical result of using the words chosen as Ayres recommends. Compare the shape of this distribution with the shape of Figs. 1 and 2. These figures were presented as evidence that teachers' marks were inaccurate. The class average is eighty-four per cent, but those pupils who spelled all of the words correctly have not been tested. Those who misspelled only one or two words probably have not been tested satisfactorily.

---

[1] The reader should not confuse scores or measures of ability with school marks. The per cent of correct spellings is a measure. The school mark is the meaning which the school attaches to that measure. The fact that both the measure and the school mark may be expressed in per cents does not make them the same.

Otis[1] presents facts from which he concludes that the most reliable measures of spelling ability are obtained by using words for which there is an average of fifty per cent of correct spellings. In support of this conclusion he points out that a list of words for which the average per cent of correct spellings was either zero per cent or one hundred per cent, would yield a measure of zero reliability. Likewise a list of words for which the average per cent of correct spellings was ten per cent or ninety per cent, would yield measures only slightly more reliable. Hence, it seems natural that the most reliable measures would be obtained by using a list for which the average per cent of correct spellings was fifty. On the other hand, some writers claim that it is not wise to have pupils spell words incorrectly. They point out that every repetition tends to fix a habit.

Ayres gives no satisfactory justification for recommending the choice of words for which an average of eighty-four per cent of correct spellings may be expected. When measuring the spelling ability of children in Springfield, Illinois, Ayres used words for which seventy per cent of correct spellings had been obtained. For the Survey of Cleveland, Ohio, the words were chosen from columns for which the average per cent of correct spellings was seventy-three. Thorndike has used words for which the per cent of correct spelling is fifty.[2] For these reasons it is probably best to choose words from columns for which the average per cent of correct spellings is approximately seventy.

---

[1] Otis, A. S., "The Reliability of Spelling Scales"; in *School and Society*, vol. 4, p. 753.

[2] Thorndike, E. L., "Means of Measuring School Achievement in Spelling"; in *Educational Administration and Supervision*, vol. 1, p. 306.

(3) **How many words to use.** Another question which must be considered in making a spelling test is the number of words it is necessary to use. In general the ability to spell one word is separate and distinct from the ability to spell any other word. Ability to spell, therefore, consists of a large number of abilities to spell specific words. This being the case it would be necessary to use all of the thousand words of Ayres's list in order to obtain a complete and accurate measure of a pupil's ability to spell the most commonly used words. However, it is possible to secure a measure which is representative of the pupil's ability to spell these words by using a smaller number of words. This is possible in just the same way that it is possible to determine the quality of a load of wheat or a vat of cream by the examination of a sample.

How many words are necessary in making a spelling test depends upon what is desired. Relying upon the theory of random sampling, Thorndike believes a small number of words is sufficient to measure the spelling achievement of a large school system. A test consisting of only ten words has been used in a number of school surveys. This number is probably sufficient for the measure of a large school system, but if it is desired to obtain a measure of the spelling ability of individual pupils, a larger number must be used. Otis[1] says that a twenty-five-word test gives a very poor measure of individual ability, and that at least one hundred words should be used, better four hundred or five hundred words. Starch recommends the use of two hundred words. Therefore, it is probably best to use as large a list of words as the time which the teacher can use for measuring spelling will permit. At least fifty words should be used if possible.

---

[1] *Loc. cit.*, pp. 679, 682.

(4) **How should the words be given.** The complaint is frequently made that pupils spell words correctly in the spelling class, but misspell the same words when writing compositions and other school exercises. One reason why this occurs is that in the spelling class the pupil has his attention fixed upon the spelling of the word and takes time to do his best. In writing a composition, his attention must be centered upon what he is writing, and thus he is able to give only partial attention to the spelling. Also he probably writes more rapidly. Hence, we may recognize two types of spelling ability: (1) the ability to spell words when one's attention is focused upon the spelling; (2) the ability to spell words when one's attention is focused upon other things and the spelling is carried on in the margin of consciousness.

The words which make up a test may be dictated to the pupils as separate words, or they may be embedded in sentences which are dictated. Furthermore, the dictation of the sentences may be timed so that the pupils are forced to write at their normal rate. In this way we are able to secure approximately the second type of spelling. Investigation has shown that the per cent of correct spellings is higher when the words are dictated separately than when they are dictated in timed sentences and the pupils are forced to write at their normal rate. According to Courtis the per cent of correct spellings is about five greater when the words are dictated in lists. Fordyce has found this difference to be between ten and fifteen per cent. The writer has found a difference of more than six per cent.

In writing letters, compositions, and the like, the spelling must be carried on in the margin of the attention because the ideas which are being expressed must occupy the focus of the attention. This is particularly true of the foundation words of the language such as we have in the Ayres list. The words of this list constitute over ninety per cent of the words we use. Hence, by using the words embedded in sentences and dictated rapidly enough to force the child to write at his normal rate, we measure the spelling ability which functions in one's every-day writing.

The rate of dictation. Pupils may be caused to write at approximately their normal rate by dictating the sentences at that rate. The Freeman's standards for rate of handwriting are as follows in terms of letters per minute: second grade, 36 letters; third grade, 48 letters; fourth grade, 56 letters; fifth grade, 65 letters; sixth grade, 72 letters; seventh grade, 80 letters; eighth grade, 90 letters. The dictation of a sentence requires some additional time, probably ten per cent. For example, in the case of the sixth grade, instead of dictating at the rate of 72 letters in one minute, 66 seconds should be allowed for words totaling 72 letters. On this basis the number of seconds to be allowed per letter for the several grades are as follows: —

Grade	*Seconds per letter*
II	1.83
III	1.38
IV	1.18
V	1.01
VI	.92
VII	.83
VIII	.73

If the sentences contain more than thirty to forty letters, they should be dictated in sections, so that the pupil's writing will not be slowed up by trying to recall what has been dictated. Furthermore, tests of rate in handwriting have shown that all pupils do not normally write at the same rate. For this reason provision must be made for those pupils who are accustomed to write more slowly than the standard rate. This can be done by having none of the test words come at the end of the sentences, and requiring all pupils to begin upon the next sentence as soon as it is dictated, even if they have not finished writing the preceding.

Summary: The discussion of the making of a spelling test may be summarized as follows:

1. The Ayres Measuring Scale for Ability in Spelling is a list of the one thousand most commonly used words of the English language. These words have been classified according to difficulty and words chosen from one column may be considered as being equally difficult. When words are taken from more than one column the inequality of difficulty must be recognized, if an accurate measure is to be secured.

2. Twenty words are probably sufficient to secure a reliable measure of the spelling ability of a class. At least fifty words should be used to secure a reliable measure of the spelling ability of individual pupils. More accurate measures will be obtained by using one hundred words. In the case of the upper grades it will be necessary to use words from more than one column. When this is done the relative difficulty of the words must be recognized to secure an accurate measure.

3. In order that the words may be difficult enough to really measure the spelling ability of all pupils, the words should be chosen from columns for which the standard per cent of correct spellings is approximately seventy. For the lower grades it is probably best to use words for which the standard per cent of correct spellings is from fifty to sixty-six. If the words are to be used in timed sentences it will probably be satisfactory to use easier words.

4. In order to secure the best measurement of spelling ability, the words should be embedded in sentences, and the sentences dictated at approximately the standard rate of handwriting for the grade. Test words should not occur at the end of the sentences.

A timed sentence spelling test. In order to illustrate the type of test described above, we reproduce the directions and a test arranged for the fourth grade. The rate of dictation of this test was determined upon the basis of measurements of the handwriting rate of six thousand Kansas school-children.

Directions for Giving a Timed-Sentence Test

1. See that the pupils are provided with two or three sheets of paper and with either pencil or pen and ink. If pencils are to be used they should be well sharpened. If pen and ink are used, good pens should be provided.
2. Make certain that all pupils understand what they are to do. It is well to give a short preliminary practice in writing from dictation if the pupils are not accustomed to it. For this purpose use some simple selection.
3. It is well not to tell the pupils that they are being tested in spelling. Under no circumstances indicate the test words by emphasis in dictating.
4. When everything is ready, say to the pupils: "I have some sentences which I want you to write as I dictate them. I am going to dictate them rather rapidly, possibly more rapidly than some of you can write. If you have not finished writing one sentence when I begin to dictate another, I want you to leave it and begin on the new sentence. If there are any words you cannot spell, you may omit them. Take time to dot your i's and cross your t's. If you have any question about what you are to do, ask it now, because you cannot ask questions after I begin to dictate."
5. Use the arrangement of the sentences which has been prepared for the grade you are testing. If a teacher has two divisions of different grades, as 5B and 6A, she must test the two divisions separately, using the test which is arranged for each grade. When the second hand of your watch is at 60 read the first sentence. When the second hand reaches the next number printed in the margin, read the second sentence. Dictate the other sentences at the time indicated. Dictate the sentences distinctly, but *do not repeat*. Be careful not to suggest the spelling of the words by unduly emphasizing certain syllables. It is advisable for the teacher to practice dictating the sentences according to the directions before attempting it with a class.

6. Stop the pupils promptly at the time indicated. *Allow no corrections or additions to be made.* Ask the pupils to turn their papers over and write their name and grade. Appoint two or three pupils to collect the papers.

A Timed-Sentence Test arranged for the Fourth Grade

Test words taken from column M of the Ayres Scale

Seconds

60 He *bought* a *railroad ticket* to the city.
41 *Collect* the *account* before Sunday.
18 *Those children* will *return* soon.
53 *Anyway* she is *ready* to go.
19 *Please omit both* names.
44 *Few change* trains here.
9 He *says* the *great office* is full.
43 *Who died* this morning?
6 The *money* for the *picture* was *paid* to us.
47 The *members* did not *understand* him.
24 *Again* he *took* the car.
46 It will *provide* an *income* in his old age.
27 The *army* had *begun* to *drill* in the park.
7 He *might begin* the *contract* next week.
47 I was *unable* to *recover* the bill.
21 I have an *extra dress* with me.
51 The *deal* was *almost* closed.
19 Did you *inform* him to *follow* the car?
56 The *past month* I was in the south.
30 *While* he *goes* home, you stay.
58 The car was *driven beside* the train.
35 I saw him *enter* the place.

When the second hand reaches 1, stop the writing.

Allow no corrections or additions to be made. Ask the pupils to turn their papers over and write their name and grade. Appoint two or three pupils to collect the papers.

Marking the test papers. The most accurate results will be obtained when the teacher marks the test papers for incorrect spellings and omissions of test words, but unless the teacher has sufficient time for this work the papers may be marked by the pupils. When this plan is followed the teacher should spell out the test words and have the pupils mark with a cross words misspelled and words omitted. (When a timed-sentence test is used no attention is given to words which are not test words.) The number of test words correctly spelled should be written at the top of each pupil's paper. This is the pupil's score. By dividing the number of words spelled correctly by the number in the test, the per cent correct is obtained.

Recording the scores. For recording the scores of a class a record sheet such as shown in Fig. 29 should be used. (This record sheet is used for a fifty-word test.) In this way the teacher obtains a statement of the number of pupils who spelled forty words correctly, the number who spelled forty-one words correctly, etc. The class score may be found by adding the scores of all of the pupils together and dividing this sum by the number of pupils. This quotient is the average. For practical purposes it is just as satisfactory and more convenient to find the median. This may be done by arranging the test papers in the order of the number of words spelled correctly, the lowest score on the bottom. The score on the middle paper is the median score.

Standards: (*1*) *The Ayres Scale.* In classifying the words of his list according to difficulty, Ayres determined the average per cent of the pupils of each grade who spelled the words correctly. Thus the words of column O were spelled correctly by 50 per cent of the third-grade pupils, 73 per cent of the fourth-grade pupils, 84 per cent of the fifth-grade pupils, 92 per cent of the sixth-grade pupils, 96 per cent of the seventh-grade pupils, and 99 per cent of the eighth-grade pupils. These per cents, which are printed at the head of each column, represent the average spelling ability of pupils in the several grades when the words are dictated in

Distribution of Pupils' Scores

Number of words spelled correctly	Number of pupils	Number of words spelled correctly	Number of pupils
50		Sub- Total	
49		24	
48		23	
47		22	
46		21	
45		20	
44		19	
43		18	
42		17	
41		16	
40		15	
39		14	
38		13	
37		12	
36		11	
35		10	
34		9	
33		8	
32		7	
31		6	
30		5	
29		4	
28		3	
27		2	
26		1	
25		0	
Sub- Total		Total	
		Median Score	

FIG. 29. SHOWING THE RECORD SHEET FOR RECORDING PUPILS' SCORES ON A SPELLING TEST OF FIFTY WORDS

lists. When the words are used in timed sentences the averages have been 5 to 15 per cent lower.

It may be seriously questioned whether the averages which Ayres gives are satisfactory standards of spelling ability for the foundation words of the language. Ayres says: "Probably the scale will have served its greatest usefulness in any locality when the school-children have mastered these one thousand words so thoroughly that the scale has become quite useless as a measuring instrument." In the past we have not had the advantage of such a list and have distributed our efforts in teaching spelling over a very much larger list of words. If we accept these one thousand words as the foundation words of our language, we should place prime emphasis upon teaching them. This being the case a satisfactory eighth-grade standard would approximate one hundred per cent for all of the words. For the preceding grades the standard would be one hundred per cent for the words of the list which the pupils had been taught. For example, the easiest nine hundred words might be used for the seventh grade, the easiest seven hundred and fifty for the sixth grade, and so forth. The use of the scale in the way Ayres suggests would seem to lead to standards of this type. The distribution of the words among the several grades and the optimum standards must be determined by experimentation.

(2) *Timed-sentence spelling tests.* A series of timed-sentence spelling tests similar to the one reproduced on page 186 was given to several thousand children in Kansas about the seventh month of the school year. The median scores are given in Table XX, and they may be used as tentative standards for timed-sentence spelling tests of the type reproduced on page 186, but it must be remembered that as we improve our teaching of spelling our standards should be raised.

TABLE XX. SHOWING MEDIAN SCORES FOR A TIMED-SENTENCE SPELLING TEST OF FIFTY WORDS

Grade	*Number of pupils tested*	*Median scores. Number of words spelled correctly*	*Per cent of words spelled correctly*	*Ayres's standards*
III..........	997	28	56	66
IV..........	1060	39	78	84
V..........	1009	33	66	73
VI..........	870	40	80	84
VII..........	826	35	70	79*
VIII..........	608	42	84	88*

* The test for the seventh and eighth grades consisted of words taken from three columns. Hence these standards are only approximate.

Causes of low class scores. As in the case of other subjects the teacher should use the results of spelling tests as a basis for planning instruction which will correct the defects that the tests reveal. A class score below standard indicates an unsatisfactory condition which may be due to one or more of the following conditions:

1. The class as a whole may be unable to spell certain words.
2. Certain pupils may be unable to spell a large number of the words of the test.
3. The errors may be rather uniformly distributed as to both words and pupils.

To determine the extent to which each condition causes the low class average, the teacher should make the following type of tabulation from the test papers. This will give a record of each pupil for each word of the test. Instead of designating the pupil by number as in this illustration their names or initials can be used at the head of the columns.

Words of the test	*Pupils*											
	1	2	3	4	5	6	7	8	9	10	11	12
catch............	c	–	–	–	c	–	–	–	–	c	–	c
black............	c	c	c	c	–	c	c	c	c	c	–	c
warm............	c	c	c	c	c	c	c	c	c	c	c	c
unless...........	c	c	c	c	c	c	c	c	c	c	–	c
clothing..........	c	–	c	–	–	c	–	–	c	c	–	–
began............	c	c	c	c	c	c	c	c	c	c	–	c

c indicates the word was correctly spelled.

Although these words are listed by Ayres in his Spelling Scale as being equally difficult for pupils in general, they are not necessarily so for particular pupils. Obviously in the class here represented "catch" and "clothing" need general emphasis, while only certain pupils need to give attention to "black," "began," and "unless." Pupil 11 has misspelled five out of six words, and hence probably is a "poor speller."

What a spelling test reveals. Such a tabulation of the results of a test is valuable because it reveals the character of the spelling ability of the class. It points out the "poor spellers." It indicates whether the class as a whole find some words difficult to spell or the misspellings are uniformly distributed. However, it must be remembered that the test contains only a limited number of words, and although the results may be accepted as indicating the nature of the conditions which exist, it cannot tell the teacher all the words for which corrective instruction must be planned. Simply to know that a pupil is below standard in ability is of little value to the teacher, because in general the ability to spell one word does not imply ability to spell another word, nor does the lack of ability to spell a given word indicate that a pupil cannot spell another word. But the fact in the above illustration that most of the class could not spell correctly "catch" and "clothing" indicates that there are several words which the class as a whole do not know how to spell. On these words class instruction is needed. When the teacher knows the type of situation with which he has to deal he should then proceed to determine the particular words which all or certain pupils need to learn to spell. At least the teacher should make a very careful diagnosis of the spelling ability of each pupil whose test score is below standard, to ascertain just what words he cannot spell of those he is expected to spell.

This is accomplished by giving the pupils below standard a test including all of the words which they are expected to be able to spell. Such a test is not for the purpose of measurement, but should be thought of as the first step in the teaching of spelling. Each pupil should be required to make from this test a list of all the words which he has spelled incorrectly. The words of this list are the ones he needs to study. It is obvious that to ask a pupil to study words which he can already spell correctly is to ask him to use his time without profit.

Class correctives. "Spelling Demons." Certain frequently used words are very frequently misspelled. Jones[1] has given us a list of one hundred words which he found misspelled most frequently in children's compositions. He calls them the "One hundred spelling demons of the English language." Nine tenths of these words are found in Jones's list for the second and third grades. Four fifths of these words are found in Ayres's list. Because these words are frequently misspelled and are among the commonly used words of the language a teacher will make no mistake in emphasizing these words in the teaching of spelling until the pupils can spell them correctly.

---

[1] See pages 176–77 for a description of Jones's study.

The One Hundred Spelling Demons of the English Language

which
their
there
separate
don't
meant
business
many
friend
some
been
since
used
always
where
women
done
hear
here
write
writing
heard
does
once
would
can't
sure
loose
lose
Wednesday
country
February
know
could
seems
Tuesday
wear
answer
two
too
ready
forty
hour
trouble
among
busy
built
color
making
dear
guess
says
having
just
doctor
whether
believe
knew
laid
tear
choose
tired
grammar
minute
any
much
beginning
blue
though
coming
early
instead
easy
through
every
they
half
break
buy
again
very
none
week
often
whole
won't
cough
piece
raise
ache
read
said
hoarse
shoes
to-night
wrote
enough
truly
sugar
straight

Class drill. Courtis[1] recommends a form of class drill which may be used when the class as a whole are learning to spell a word:

> The word to be learned is pronounced by the teacher and class together and then written letter by letter as it is spelled aloud. This is repeated five or six times in rapid succession. The rate of writing should be slow at first, then faster and faster (like a college yell), until at the sixth repetition only the most rapid writers are able to keep up with the class.

Spelling games. In the manual referred to above, Courtis describes the following games which may be used for providing drill upon spelling. They are particularly helpful when a stronger motive is needed. Each of these games provides for dividing the pupils in a room into two teams or groups and for keeping scores for a week or a month:

---

[1] Courtis, S. A., *Teaching Spelling by Plays and Games* (82 Eliot Street, Detroit, Michigan), p. 8. This is an excellent manual for teachers. It contains explicit directions for a number of spelling games.

1. Syllable game.
2. Jumbled-letter game.
3. Initial game.
4. Rhyming game.
5. Derivative game.
6. Definition game.
7. Linked-word game.
8. Missing-word game.
9. Composition game.

Individual correctives. Types of misspellings. A pupil's spelling difficulty is not completely diagnosed when the words he does not spell correctly are located. Errors in spelling are seldom if ever distributed uniformly among the several letters composing the word. Neither does it appear that there is much uniformity in the location of errors in different words. Certain words are misspelled in only a few ways, while other words are misspelled in many ways. Certain misspellings occur frequently, while others seldom occur. In Table XXI the misspellings of certain words found in the papers of eighty seventh-grade pupils are given, together with the frequency of each. The words were taken from column S of the Ayres Scale. Where no number follows the word that type of misspelling of the word occurred but once.[1]

Causes of some misspellings. A study of Table XXI shows that certain forms of misspelling occur more frequently than others, and that most of the misspellings may be attributed to certain specific causes. Forms of misspell-

---

[1] See also Sears, J. B., *Spelling Efficiency in the Oakland Schools.* Board of Education Bulletin, Oakland, California, p. 51.

TABLE XXI. THE MISSPELLING OF EIGHTY SEVENTH-GRADE PUPILS ON A COLUMN SPELLING TEST

- I. affair
  - affere
  - affire
  - afair (2)
  - affaired
  - affer
- II. assist
  - assit (3)
  - aisst (2)
  - ascist
  - assest
  - assaist
  - asscest
  - assiast
  - acsist
  - acist (2)
  - accisted
  - assantant
  - assised
  - accessese
  - accest
  - astist
  - assis
  - assite
- III. certain
  - certian (7)
  - serten
  - sertain
  - certin
  - secrtain
- IV. difference
  - differance (10)
  - diffierence
- V. examination
  - examation (10)
  - examition
  - examnition
  - excamation
  - excanitions
  - examanation (3)
- VI. government
  - goverment (9)
  - govament
  - governement
  - gorvement
- VII. improvement
  - improvment (7)
  - impovement
- VIII. investigate
  - investagate (3)
  - envesigatige
  - investiage
- IX. marriage
  - marrage (5)
  - marage
  - merriage
- X. mention
  - mension (8)
  - mensioned
  - meantion (2)
  - menchion
- XI. motion
  - moshen
  - moticem
  - motation
  - montion
- XII. neither
  - neather (6)
  - nether
  - niether (2)
  - nieghter
- XIII. opinion
  - oppinion (5)
  - opinon (2)
  - opinton
  - oppoinen
  - oppinum
  - oppenion (2)
  - opion (3)
  - oponion (2)
  - oppion (2)
  - opinnion
  - opoin (2)
  - opionion
- XIV. particular
  - particuliar
  - particuler
  - partictuler,
  - perticular (8)
  - particlar
  - pertucular
  - partcular
  - partiular
  - partular
  - paticular
  - perciluar
  - pectuliar
  - pecticular
  - pertictural
  - patuclure
  - peculiar
  - pectulair
- XV. possible
  - possable (4)
  - posible
  - posable
  - posiable (2)
  - possiable (5)
  - posobile
  - possibbe
  - posiple
- XVI. serious
  - cyreaus
  - cerrious
  - scerious
  - serrious (2)
  - cerious
  - sereaus
- XVII. stopped
  - stoped (13)
  - stopts
  - stocted
  - stop

ing such as "partiular," "partuler," "opinon," "impovement," "possibbe," are probably due to carelessness or accident, "a slip of the pen." Relatively few of the misspellings in this table may be assigned to this cause. Errors of this type probably cannot be entirely eliminated from uncorrected manuscript. However, drill will reduce the number of such errors to a satisfactory minimum.[1]

An important source of error is mispronunciation of the word by the pupil. He may have acquired this from the teacher, but more likely from those with whom he associates outside of school. Or it may have been acquired from lack of attention to the form of the word. Such misspellings as the following are probably caused by mispronunciation: "perticular," "particlar," "investagate," "goverment," "examation."

A very striking instance of this type of spelling error and its cause came to the attention of the writer a few years ago. A man who had taught geometry for a number of years used the word "frustum" in a manuscript, spelling it "frustrum" which agreed with his pronunciation of the word. This manuscript was read by a number of well-known mathematicians who read it critically. Only two noted the misspelling of the word, and one mathematician, who took much pride in his ability to spell correctly and who was the author of several textbooks, admitted that he had always pronounced and spelled the word "frustrum."

Other errors listed in Table XXI are due to certain phonic irregularities of the English language, for example, certain misspellings of "assist," "certain," "affair," "marriage," "motion," "neither," and "serious." Such errors occur more frequently in connection with vowels than with consonants. Still other errors, such as "stoped," and "improvement," are due to certain doubled or silent letters.

The length of words and the position of the letters are responsible for some errors. In general there is a close agreement between the number of letters in a word and its relative difficulty. The longer the word, the more difficult it is to spell. In Table XXI it is obvious that the errors are not uniformly distributed among the several letters of a word. For example, consider "examination," the fifth word in the table. The letters *e-x* and *t-i-o-n* were given correctly in every case. The first *a* also occurs. In every case except one, the letter *m* is given. Fifteen out of the seventeen errors occur in connection with three letters, *i-n-a*. The explanation of this condition, which is typical, is that correct spelling "depends mainly upon a correct visual or audile image coördinated with the correct motor control."[1] Some letters are more conspicuous than others in the form of the printed, or written, word and also in the sound of the spoken word. In general, the letters occupying the initial positions are remembered best for this reason.

---

[1] Errors of this type have been called "lapses." See Hollingworth, Leta S., *The Psychology of Special Disability in Spelling*, Teachers College, Columbia University, Contributions to Education, no. 88, p. 38, for a complete statement of types.

Some words are spelled incorrectly because the pupil has not learned any spelling, correct or incorrect, for the word. In such cases if the pupil is asked to spell the same word several times, different spellings will be given. Hollingworth gives an illustration of this type. One pupil misspelled "saucer" in seven different ways in nine successive writings of the word: "s-a-u-e-c," "s-u-s-s-e," "s-u-c-c-e-r," "s-u-c-c-e-r-e," "s-u-r-r-e-s," "s-u-s-s-e-r," "s-u-c-e," "s-u-s-s-e-r," "s-u-c-c-e-r." Other words are misspelled because the pupil has learned an incorrect spelling. This was the case in the misspelling of "frustum" described above.

Still another cause of misspelling is a lack of the knowledge of the meaning of the word. Hollingworth states:[2]

> On the basis of these data we conclude that *knowledge of meaning is probably in and of itself an important determinant of error in*

---

[1] Kallom, Arthur W., "Some Causes of Misspellings"; in *Journal of Educational Psychology*, vol. 8, p. 395.

[2] *Psychology of Special Disability in Spelling*, p. 57.

spelling; that children will produce about sixty-six and two-thirds per cent more of misspellings in writing words of the meanings of which they are ignorant or uncertain, than they will produce in writing words the meaning of which they know.

Teaching the pupil to correct his errors in spelling. Spelling consists in forming correct and fixed associations "between the successive letters of a word and between the word thus spelled and the meaning."[1] The laws governing the formation of fixed associations are those of habit formation. The first step in habit formation is to get the attention of the child focused upon the associations to be formed. The second step is to secure sufficient repetition. Repetition of the associations is secured both through drill and through using the word in written expression. The pupil must give attention to the repetitions of the associations in order to insure that wrong associations will not be made.

The causes of misspelling given above suggest certain correctives. If the error is due to an incorrect pronunciation of the word, the pupil should be taught the correct pronunciation. The phonic irregularities of words should be emphasized. In the case of long words the pupil's attention should be directed to the letters in the middle of the word. The meaning of the word should be connected with the pupil's experience. This does not mean merely requiring him to use it in a sentence.

As in the case of other school subjects motive is an important factor in learning to spell. A strong motive can be secured by the use of standardized spelling tests. Definite standards should be set and at intervals careful tests should be made. Charts showing the scores of the individual pupils as well as the class score in comparison with the standard will be helpful.

---

[1] Freeman, F. N., *Psychology of the Common Branches*, p. 115.

Numerous experiments have shown that pupils can spell correctly a large per cent of the words in the lists in spellers before they have studied them. Because of this fact the assignment of the spelling lesson should include the dictation of the words to the pupils so that each might know what words he needed to study. The teacher would also learn what words he should emphasize in his instruction.

Some writers state that a pupil should not be permitted to spell a word incorrectly when it can be avoided, and for this reason pupils should learn to spell words correctly before they are required to write them. Just how important it is to do this we do not know. In certain cases it appears that a child or an adult learns to spell certain words correctly by having his attention directed to his errors. The fact of his error serves to direct his attention to learning to spell the word correctly. Those who believe that evil effects will come from having pupils write words which they cannot spell correctly, may direct them to omit those words which they think they cannot spell correctly.

The dictation of the words in assigning the spelling lesson, together with the detailed testing of the pupils as suggested on page 191, reveals to the teacher the words upon which he must exercise his ability as a teacher of spelling. It also reveals to him the pupils to whom instruction should be directed in the case of each word. Particular methods and devices by means of which the laws of habit formation may be fulfilled are described in books which deal with the teaching of spelling.[1]

A device for focusing attention upon the difficult portion of a word. In teaching the spelling of a word the child's attention should be directed to the crucial associations. If the word is one like "government," his attention should be called to the correct pronunciation. If it is such a word as "their," his attention should be called to the use of the word. To eliminate spelling errors a pupil's attention should be called to his particular error and he should be helped to remove the cause. If the cause is mispronunciation, see that he learns to pronounce the word correctly. If the error is due to a confusion of letters, the pupil should be given some device to prevent this confusion. The following is a device which may be used for especially difficult words:

---

[1] A very good chapter (VI) will be found in Freeman's *Psychology of the Common Branches*. See also Cook and O'Shea, *The Child and his Spelling*.

Par-tic-u-lar

I frequently misspell .... in writing compositions but now I am going to learn to spell it correctly. My teacher tells me that I do not look at the .... syllables and letters closely enough. I am going to do it now with .... care. I see that the word has .... syllables. The first syllable is .... The vowel of this syllable is ...., the first letter of the alphabet. The last syllable is .... and the vowel is also .... The word contains .... letters, the other vowels are .... and .... Now that I have looked at the word carefully I am going to be very .... in spelling it. I am also going to be .... in pronouncing it. I am going to remember that the vowel in the first syllable and in the last syllable is an .... I am not going to pronounce those syllables as if the vowel were *e* instead of .... I am going to be very .... about both spelling and pronouncing this word. I want it to be correct in every..........

This device is used by providing the pupil who needs instruction with a printed or typewritten copy. The pupil is required to fill in the blank spaces correctly. This is repeated until the correct associations are fixed.

Drill for making associations automatic. Getting the pupil to spell a word correctly is only the first step. There must be attentive repetitions of the correct associations until they have become automatic. In this respect spelling is similar to arithmetic. In the teaching of the operations of arithmetic, drill occupies a prominent place, but in the case of spelling our teaching has been confined primarily to testing pupils. Requiring pupils to write each misspelled word ten or twenty times is an effort to provide practice. Such practice is unsatisfactory. After the first writing of the word the pupil probably copies. Hence the repetitions are not attentive.

Practice upon words which are misspelled by a majority of the pupils can be secured by having them recur in the spelling lesson from day to day. This plan provides the same drill for all pupils regardless of whether they misspell the word or not. In this respect it is unsatisfactory. The pupils may be required to write material which the teacher dictates. When carrying on this kind of practice, the teacher dictates as rapidly as the pupils can write, or better, calculates the number of seconds required to write each sentence, or part of sentence, as was done for the timed-sentence spelling test. (See page 186.) This can easily be done by using the rates given on page 183.

"Developing a spelling consciousness." The following device serves to direct the pupil to see his errors in a wholesome way. It has yielded very gratifying results in the Training School of the Kansas State Normal: [1]

When the spelling sentences or lists have been written, each pupil is required (1) to mark each word, the spelling of which he doubts; (2) as far as possible he is encouraged to test the validity of his doubts by known means outside of the dictionary, finally checking up all doubted words by using the dictionary; and (3) he then writes all of the misspelled words, which he has thus detected, correctly spelled in separate lists; (4) at this point the pupils' papers are exchanged, the teacher spelling all words and the pupils marking those found to be misspelled on the papers; and finally (5), when the papers are returned to their owners the additional misspelled words discovered should be added to their individual lists.

---

[1] Lull, Herbert G., "A Plan for Developing a Spelling Consciousness"; in *Elementary School Journal*, vol. 17, p. 355.

The pupil's spelling is scored by the teacher on the basis of the correctness of his doubts as well as upon the number of words spelled correctly. In the absence of a scientific determination of the relative significance of *spelling of words correctly* and *doubting correctly*, the same value is assigned to each. The pupils are scored both for doubting words spelled correctly, and for not doubting words spelled incorrectly.

QUESTIONS AND TOPICS FOR STUDY

1. Measure the spelling ability of the pupils of your class by means of a timed-sentence test and then dictate the test words as separate words. Compare the two sets of scores.
2. Teachers frequently tell with pride that all but two or three of their pupils make a "grade of 100" on a certain test. Should the fact be a cause for a feeling of satisfaction? Were the pupils really tested?
3. Dictate the words for the next spelling lesson before the pupils have studied them. Have each pupil make a list of the words which he misspells and also of the particular misspellings which he has used. Direct the pupils to base their study upon these lists.
4. Construct a series of "timed-sentence spelling tests" for the elementary school, using suitable words from the Ayres Scale.
5. Why does a test of easy words fail to give a measure of spelling ability?
6. Why must the relative difficulty of the words of a test be known if accurate measures are desired?
7. Make a study of the ways in which your pupils misspell words. Also ascertain the causes for these misspellings.
8. How can you use this information in making your teaching of spelling more effective?

CHAPTER VIII

THE MEASUREMENT OF ABILITY IN HANDWRITING

THE measurement of ability in handwriting involves (1) the measurement of the rate of writing and (2) the quality. The rate is measured by having the pupil write under specified conditions for a convenient number of minutes and counting the number of letters written per minute. The measurement of quality is accomplished by securing a sample of the pupil's handwriting and determining the specimen on a handwriting scale which is equivalent to it in quality. The quality may also be measured by means of a score card.

The measurement of the rate of handwriting. In measuring the rate of handwriting certain points must be kept in mind.

(1) The teacher should see that all pupils are provided with good pen-points, ink, and paper unless they use pencils, in which case there should be a sufficient supply of well-sharpened pencils. All pupils should be supplied with two or three sheets of suitable writing-paper.

(2) The pupils should be asked to write a sentence or a paragraph which they have memorized. To guard against lapses of memory, the pupils should be asked to repeat in concert the selection to be used. If convenient it is well to provide each pupil with a printed or typewritten copy of the selection. When this cannot be done, the selection may be written on the blackboard where all can see it. The selection should contain no words which the pupils cannot spell readily. It is well to have them practice writing the more difficult words before the test is begun. Do not use material which the pupils must compose as they write, for

this would be worthless in testing. The rate of writing unfamiliar material from a printed copy will vary with the pupil's rate of reading and so will not give a true measure of his rate. Dictated material should be used only when the teacher wishes to control the rate, not when the rate is to be measured.

(3) The teacher must be provided with a watch which has a second-hand or with a stop-watch. A two- or three-minute period should be allowed and the teacher should exercise care to make this period exact.

(4) Pupils probably have two or more rates of writing, one for the penmanship class when they are doing their best and another for writing exercises in history, language, or other school subjects. The way in which the teacher gives the directions to the class will influence their rate. If he tells the pupils or even suggests that they are expected to show how well they can write, the rate will probably be low. On the other hand, if the pupils are given the idea that the rate is most important, they will write more rapidly than they are accustomed to do. It is, therefore, important that a teacher follow directions which have been prepared for securing samples of pupils' handwriting. We give below a set of directions which have been widely used.

Directions for Obtaining Samples

When children have paper and pencil proceed thus: Read aloud a stanza — four lines of the poem, "Mary had a little lamb," etc., which is printed below. If you are using these directions for the first time use the first stanza; if the second time, use the second stanza; if the third time, use the third stanza. Have the children recite this stanza aloud in unison until you are sure they all know it. Then ask them to write it once. Collect these copies and destroy them. *Do not tell the children they are to be tested in any way.* Next instruct the children as follows:

"Write the stanza of the poem which you have learned. Write it just as you would in a composition or in an ordinary school exercise. If you finish the stanza, write it over again, and keep on writing until I tell you to stop. Write on only one side of the paper. When you fill one page use another. We must start together and stop together. Lay your papers on your desk in position. Have pen and ink ready. When I say 'Get ready,' ink your pen and place your hand in position to write, but do not begin to write until I say 'Start.' Then all begin at once. When I say 'Stop' I want you all to stop at once and raise your hands so that I can see that you have stopped."

Now take your watch in hand and when the second-hand reaches the *55* second mark say, "Get ready." Exactly at the 60 second mark say, "Start." At the end of three minutes call out, "Stop, hands up." Be sure to allow exactly three minutes. Have each child write name and age on the *back* of the paper. Collect the samples at once and put them together.

I

5 10 15
Mary had a little lamb,
20 25 30 35 40
Its fleece was white as snow;
45 50 55 60 65
And everywhere that Mary went
70 75 80 84
The lamb was sure to go.

II

5 10 15 20 25
He followed her to school one day;
30 35 40 45
That was against the rule;
50 55 60 65 70 75
It made the children laugh and play
80 85 90 95 97
To see the lamb in school.

III

5 10 15 20 25
And so the teacher turned him out,
30 35 40 45
But still he lingered near,
50 55 60 65 70
And waited patiently about
75 80 85 89
Till Mary did appear.

Directions for securing samples for the "Gettysburg Edition" of the Ayres Scale. With the "Gettysburg Edition" of his handwriting scale Ayres gives directions which should be followed when that scale is used. The directions are not entirely complete and should be supplemented by the last two paragraphs of the directions given above.

To secure samples of handwriting the teacher should write on the board the first three sentences of Lincoln's *Gettysburg Address* and have the pupils read and copy until familiar with it. They should then copy it, beginning at a given signal and writing for precisely two minutes. They should write in ink on ruled paper. The copy with the count of the letters is as follows:

Four 4 score 9 and 12 seven 17 years 22 ago 25 our 28 fathers 35 brought 42 forth 47 upon 51 this 55 continent 64 a 65 new 68 nation 74 conceived 83 in 85 liberty 92 and 95 dedicated 104 to 106 the 109 proposition 120 that 124 all 127 men 130 are 133 created 140 equal 145. Now 148 we 150 are 153 engaged 160 in 162 a 163 great 168 civil 173 war 176 testing 183 whether 190 that 194 nation 200 or 202 any 205 nation 211 so 213 conceived 222 and 225 so 227 dedicated 236 can 239 long 243 endure 249. We 251 are 254 met 257 on 259 a 260 great 265 battlefield 276 of 278 that 282 war 285.

Other selections which have been used. Different investigators have required pupils to write different material. Several have used the first line or the first stanza of the poem "Mary had a little lamb," which is reproduced above. "Sing a song of sixpence" has been used. Other sentences which have furnished copy are: "Jolly kings bring gifts while happy maids dance." "A quick brown fox jumps over the lazy dog." [1] "Then the carelessly dressed gentleman stepped lightly into Warren's carriage and held out a small card. John vanished behind the bushes and the carriage moved along down the driveway." [2]

---

[1] This sentence was used in securing specimens for the Freeman Scale. It contains all of the letters of the alphabet.

[2] These sentences were used in securing the specimens for the Thorndike Scale.

In the Cleveland Survey the first three sentences of Lincoln's *Gettysburg Address* were written, and Ayres has used this selection in the "Gettysburg Edition" of his scale. In several school surveys the pupils were allowed to write any familiar stanza of a poem. The chief principles to bear in mind in selecting materials are: (1) to use material in the lower grades which will not furnish difficulties in spelling and remembering; and (2) to use material which will be uniform in all classes which are to be compared.

Marking the papers for rate of handwriting. Time can be saved by making use of the numbered selections given above. Divide the total number of letters written by the number of minutes allowed. The quotient is the number of letters per minute. This is the pupil's rate score and should be written in the upper right-hand corner of his paper.

Measuring the quality of handwriting by means of scales. The "quality" of a sample of handwriting may be measured by means of a "handwriting scale" which consists of a number of specimens of handwriting arranged in order of "quality." The process of measurement simply consists of moving the sample which is being measured along the scale until a specimen of the scale is found which "matches" it in "quality." The process is much like "matching" a sample of dress material or ribbon. Skill in this "matching" or use of the scale comes with practice and it is recommended that a teacher prepare himself by at least a short period of training.

Handwriting scales. The scales in most general use are the ones constructed by Thorndike[1] and by Ayres.[2]

---

[1] Thorndike, E. L., "Handwriting"; in *Teachers College Record* (March, 1910), vol. 2, no. 5. The scale may be purchased from the Bureau of Publications, Teachers College, Columbia University, New York City.

[2] Ayres, L. P., *A Scale for Measuring the Handwriting of School-Children.* (Russell Sage Foundation, Bulletin 113.) Ayres has also constructed an adult scale and the "Gettysburg Edition." In this book the term "Ayres's Scale" refers to the "Gettysburg Edition" unless otherwise noted.

Thorndike constructed his scale on the basis of three characteristics — beauty, legibility, and general merit. The degree of these characteristics represented in the specimens of the scale was determined by the consensus of opinion of competent judges. Ayres constructed his scale on the basis of legibility alone. He defined legibility in terms of ease of reading. That specimen was defined as most legible which was read most easily. The numerical values of the specimens of the Thorndike Scale range from 4 to 18, one or more specimens being given for each degree of quality.

Ayres's Scale, "Three-Slant Edition," consists of three types of specimens, vertical, semi-slant, and full slant. Each of these three types is represented by eight degrees of quality to which are assigned the numerical values 20, 30, 40, up to 90. In using this scale it must be remembered that these values are not the same as the per cents used in reporting "grades."

Ayres[1] later devised a scale from specimens of handwriting written by adults. Trained judges used the "Three-Slant Edition" in selecting the specimens and in determining their values. This "Adult Scale" is similar to the "Three-Slant Edition" in its general plan. Very recently (1917) Ayres devised a third scale, the "Gettysburg Edition." This scale differs from the others in the following particulars: It has one specimen for each step. The specimens are written on ruled paper. The copy is the same for all specimens. In addition to the specimens of the scale, this edition has directions for securing specimens from a class and for scoring these specimens. It also furnishes standards for rate and quality of handwriting for the grades above the fourth. Ayres asserts that the purpose of these changes is "to increase the reliability of measurements of handwriting." A recent investigation[1] shows that measurements made by this scale are more reliable than when made by the "Three-Slant Edition."

---

[1] Ayres, L. P., *A Scale for Measuring the Handwriting of Adults.* (Russell Sage Foundation, Bulletin E 138.)

Following discussion based on the "Gettysburg Edition" of Ayres's Scale. Because more reliable measurements may be obtained by using the "Gettysburg Edition" of Ayres's Scale, we shall base the following discussion upon it. It will, however, be an easy matter for any one to adapt it to any other scale. In order to understand the following pages the teacher should have a copy of this scale. See the Appendix for price list and directions for securing a sample package of tests.

Training in the use of a handwriting scale. The accuracy of a teacher's measurements of quality handwriting depends upon the method he uses and upon his training in the use of that method. When a teacher is using a handwriting scale for the first time the following preliminary exercise is recommended:

> Select ten samples at random. Number these samples and place their numbers on a blank sheet of paper. Now take the first sample and rate it thus: Place the Ayres Scale on a table in full view and in a good light. Place the sample directly under the scale division marked 20 and move it along toward 90, comparing it with each division. Decide which division of the scale it resembles most in quality. "Disregard differences in *style*, but try to find on the scale the *quality* corresponding with that of the sample being scored." Then place it under the scale division marked 90 and work back toward 20 as before. Decide again which division it resembles most in quality. If your two judgments agree, mark the rating on the blank paper opposite the numeral 1. If the two judgments do not agree, compare the sample again with the two divisions of the scale and determine which it most nearly resembles. Proceed to rate the other samples in this manner, keeping the record for

---

[1] Breed, F. S., "The Comparative Accuracy of the Ayres Handwriting Scale, Gettysburg Edition"; in *Elementary School Journal* (February, 1918), vol. 18, p. 458.

each. When you have finished the ten samples, lay this record aside, out of sight. Rate the ten samples a second time, again keeping the records and again laying the records aside. Do this a third time, and when you have finished, compare your three ratings for each of the ten samples. If the three ratings for any one sample vary more than ten, satisfy yourself as to which rating is the correct one, by comparing it with the scale again.

When convenient it is better to use samples whose correct rating is known. A set of fifty such samples may be obtained from the Bureau of Publications, Teachers College, Columbia University, New York City. They are rated in terms of Thorndike's Scale, but these scores can be changed to Ayres's Scale by multiplying by 6.7 and subtracting 20 from the product. The remainder is the true quality of the sample on the Ayres Scale.

Method of using the scale. For using the "Gettysburg Edition" Ayres gives the following directions:

> *To score samples* slide each specimen along the scale until a writing of the same quality is found. The number at the top of the scale above this shows the value of the writing being measured. Disregard differences in *style*, but try to find on the scale the *quality* corresponding with that of the sample being scored. With practice the scorer will develop the ability to recognize qualities more rapidly and with increasing accuracy. If the scoring is done twice, the results will be considerably more accurate than if done only once. The procedure may be as follows: Score samples and distribute them in piles with all the 20's in one pile, all the 30's in another, and so on. Mark these values on the backs of the papers, then shuffle the samples and score them a second time. Finally make careful decisions to overcome any disagreements in the two scorings.

Whenever three or more persons can work together in scoring specimens the results may be expected to be more satisfactory than those secured by independent work. All the members of the group should examine the specimen of writing and confer concerning the rating it should receive.

A majority of the group must agree before a score is assigned to the specimen.

A method which will require more time, but one which will secure more accurate results than the methods described above, is one in which a group of three or more persons score the specimens independently, using the sorting method. Then the scores assigned by all of the judges to a specimen are averaged and the result taken as the true score for that specimen. The accuracy of the resulting scores will increase with the size of the group of judges.

Recording scores. After the samples are rated the teacher must be careful that his papers are grouped correctly by classes. If he has but one grade of pupils, say fifth grade, or two divisions of the same grade, say fifth A and fifth B, then his papers may be all grouped together and but one distribution made. If, however, he has parts of two or more grades, say part fifth and part sixth, he must fill out a separate record sheet for each division. A convenient form of a class record sheet is shown in Fig. 30.

Sort the papers from one class on the basis of quality. (For instance, put into one pile all those papers having a quality of 90, into another put all the 80's, into another all the 70's, and so on.) Then, one pile at a time, re-sort the papers in each of these piles on the basis of their score for rate, placing together those papers whose rates are 30 to 39, 40 to 49, 50 to 59, etc. (For example, if there were ten papers of quality 60, whose rates were 50, 53, 55, 62, 62, 64, 69, 72, 77, 80, the first three would be piled together, the next four would form a second pile, the next two a third pile, and the last one would be placed by itself.) Next count the number of papers in each of these piles and record the numbers in the proper vertical column of the table. (In our illustration this is the column under 60. There are three papers in the pile whose rates are between 50 and 59. Place a figure 3 in the 60 column and directly opposite the numerals 50 to 59. There are four papers in the pile whose rates are 60 to 69. Hence, a figure 4 is to be placed in the 60 column and opposite the numerals 60 to 69.) Each of the other piles is to be treated in the same way.

When all the scores have been entered, find the sum of the figures in each vertical column and in each horizontal row. If your records have been accurately made, the sum of the horizontal totals will just equal the sum of the vertical totals. Save all specimens for future use.

Computing class scores. The medians of the distributions (rate and quality) are used to designate the general standing of a class. The method of calculating the median, described on page 103, is used. It is necessary to remember that in the record sheet shown on page 213, the width of the intervals is 10, the same as in the accuracy distributions for arithmetic. When there are fewer than fifteen pupils in a class it is not wise to attach much importance to the medians. The distributions and individual scores are more significant.

Measurement for diagnosis. The quality of a sample of handwriting is a complex product. It depends upon several characteristics of the handwriting, such as the uniformity of slant, uniformity of alignment, letter formation, and spacing. There are available two instruments for diagnostic measurement:

1. Freeman's[1] *Scale* which differs from the other scales in an important respect. It is in reality five scales, one for each of the following characteristics of handwriting: uniformity of slant, uniformity of alignment, quality of line, letter formation, and spacing. These five scales are now printed

---

[1] Freeman, F. N., *The Teaching of Handwriting.* (Houghton Mifflin Company, 1915.) Also, "An Analytical Scale for the Judging of Handwriting"; in *Elementary School Journal* (April, 1915), vol. 15, p. 432. A copy of the scale can be obtained from Houghton Mifflin Company, Boston. Price 25 cents.

DISTRIBUTION OF PUPILS' SCORES

Number of letters written in one minute	*Quality*								*Total for rate*
	20	30	40	50	60	70	80	90	
Below 10									
10 to 19									
20 to 29									
30 to 39									
40 to 49									
50 to 59									
60 to 69									
70 to 79									
80 to 89									
90 to 99									
100 to 109									
110 to 119									
120 to 129									
130 to 139									
140 to 149									
Over 150									
Total for quality									

Approximate Class Medians : Quality..............................Rate (Letters per min.)............
True Medians : Quality......................................Rate (Letters per min.)............

FIG. 30. SHOWING FORM OF CLASS RECORD SHEET FOR RECORDING SCORES IN HANDWRITING

on one sheet of paper or chart, and each scale is called a division.

The first of the five divisions of the Freeman Scale represents three degrees of uniformity of slant. In using this division, as in using the next division, judgments will be made more easily if a slant and alignment gauge is used.[1] The second division represents uniformity of alignment. The user must be careful to note that letters which are close together show deviations in alignment more prominently than letters written farther apart.

The third division shows the quality of line or stroke. A reading-glass will aid in judging with this division. The fourth division is intended to measure letter formation. Freeman describes eight illegible forms of letters which should be counted as errors. Two principles should control here: (1) whatever slant or type of script the pupil may use, consistency to that choice should be maintained; and (2) no letter should vary from its recognized form so much as to be easily mistaken for another letter. The fifth division shows different kinds of spacing. Letters may be crowded or spread too far apart. The same applies to words.

In each division the three degrees of excellence are given scores of 1, 3, and 5 respectively. The intermediate values of 2 and 4 may also be used. If the old edition of the scale is used, the scores assigned to the specimens of letter formation are 2, 6, and 10. Freeman suggests that the specimens be scored by using the score for letter formation as placed on the new edition of the chart, and then doubling these scores in making up the total score.

---

[1] Freeman, F. N., *The Teaching of Handwriting*, p. 151. The slant gauge consists of three rows of parallel lines. The lines in one row are vertical and in each of the other rows the lines are set at a uniform slant. The alignment gauge consists of one straight line four or five inches long. These lines may be drawn on transparent paper and placed over a specimen of handwriting to assist in determining the deviations from uniformity in slant and alignment.

Using the Freeman Scale. This scale may be used for measuring specimens from all members of a class, but frequently it is used to measure specimens written by those ranking conspicuously below the average ability or below the standard ability for the class. This needy group of pupils may be selected by the teacher's unaided judgment, but preferably by the use of the Thorndike or Ayres Scale.

Freeman [1] has recently issued the following suggestion for using his scale:

> The specimen to be judged is graded according to each category separately and given the rank of the specimen in the chart with which it most nearly corresponds in each case. The total rank is calculated by summing up the five individual ranks. Thus, if letter formation is given double value, the lowest possible rank is 6 and the highest possible rank is 30 (5 + 5 + 5 + 10 + 5), and the range is 24.
>
> Several precautions are to be observed in making the judgments. The value of the method rests upon the fact that different features of the writing are singled out, one at a time, and graded by being given a rank in one of only three steps. The difference between the steps are marked, and the ease of placing a specimen should be correspondingly easy.
>
> This method implies, however, that
>
> (1) The attention is fixed on only one characteristic at a time.
> (2) The judgment on one point be not allowed to influence the judgment on the other point.
> (3) The same fault be counted only once.
> (4) General impressions be disregarded.

The scores secured by means of the Freeman Scale should be saved to furnish a means of evaluating the results secured from instruction. The scores may be recorded on the specimen, or, better, on an individual record card, such as shown in Fig. 31. The latter will be more convenient when the

---

[1] Freeman, F. N., *Experimental Education,* p. 86. (Houghton Mifflin Company, 1916.)

Pupil's Name.....................City..................

	First trial *Date.......*	*Second trial* *Date.......*	*Third trial* *Date.......*	*Fourth trial* *Date.......*	*Teacher.........................* *Grade............ Age..................* *Building.........................*
Chart I (Slant)					
Chart II (Alignment)					
Chart III (Quality of line)					
Chart IV (Letter formation)					
Chart V (Spacing)					
Total (value on Freeman Scale)					
Quality (value on Ayres Scale)					
Speed (Letters per minute)					

Fig. 31. Individual Record Card, Freeman Scale.

teacher wishes to examine a series of scores recorded at intervals over a term of several months.

2. Gray's Score Card for detailed analysis. The score card represents another attack upon the problem of measurement. It requires that the essential elements of handwriting be selected and each assigned a value. The score card devised by Gray[1] weights the value of each of the essential elements of handwriting so that the highest value which can be assigned to slant is 5, while spacing of letters may receive 18, neatness, 13, etc. (See Fig. 32.) The use of this score card by teachers in their grading of handwriting would undoubtedly tend to direct their attention to the individual needs of the pupils. So far there is no evidence to show that its use will result in more accurate measures than the use of any one of the scales. Some claim that

---

[1] Gray, C. Truman, *A Score Card for the Measurement of Handwriting.* (Bulletin of the University of Texas, no. 37, July, 1915.)

Pupil........................ *Age*............... *Date*........................

Grade................................. *School*..................................

Sample Number................ *Teacher*..

Sample	*Perfect score*	*Score*													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Heaviness..........................	3	..	..	..	..	..	..	..	..	..	..	..	..	..	..
2. Slant .... Uniformity Mixed	5	..	..	..	..	..	..	..	..	..	..	..	..	..	..
3. Size .... Uniformity Too large Too small	7	..	..	..	..	..	..	..	..	..	..	..	..	..	..
4. Alignment..........................	8	..	..	..	..	..	..	..	..	..	..	..	..	..	..
5. Spacing of lines .... Uniformity Too close Too far apart	9	..	..	..	..	..	..	..	..	..	..	..	..	..	..
6. Spacing of words Uniformity Too close Too far apart	11	..	..	..	..	..	..	..	..	..	..	..	..	..	..
7. Spacing of letters...................... Uniformity Too close Too far apart	18	..	..	..	..	..	..	..	..	..	..	..	..	..	..
8. Neatness........................... Blotches Carelessness	13	..	..	..	..	..	..	..	..	..	..	..	..	..	..
9. Formation of letters..................	(26)														
General form	8	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Smoothness.........................	6	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Letters not closed....................	5	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Parts omitted........................	5	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Parts added..........................	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Total Score....................	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Fig. 32. Standard Score Card for measuring Handwriting. (*Devised by C. T. Gray.*)

the elements of handwriting have not been correctly evaluated. However, it has the advantage that its use trains the user in the analysis of handwriting. Gray well defends the device by saying that agriculturists have long used such score cards to secure very satisfactory and accurate results in judging grain and live-stock.

In using Gray's Score Card and the Freeman Scale, measures of each of the several factors concerned in a pupil's handwriting are secured. A record of successive measurements will show just what abilities have not been sufficiently improved. These abilities will then be the points of attack for the teacher and pupil in their subsequent work. For example, a record as shown on the Gray Score Card might indicate that a pupil's handwriting was suffering chiefly because of poor letter formation. A closer inspection would show that letter formation was very often defective in two items, letters not closed and parts omitted. Such diagnosis reveals a definite problem for the teacher.

Use of the score card. The score card (see page 217) may be used for a pupil, or a class. If it is used for a pupil, the numerals along the top may be taken to indicate weeks, months, or other intervals. In the column under the numeral 1 the first scores of a pupil's handwriting should be entered. A month later a second series of scores should be entered in the column headed by the numeral 2. The next month another series of scores should be entered under numeral 3, and so on. At the close of a term there will appear a very useful record of the child's experience in the learning of handwriting. This use of the score card Gray calls a clinical study.

If the card is used for a class, the numerals at the head of the columns stand for the specimens written by the several pupils of the class. The totals at the bottom will furnish an interesting comparison of the ability of the pupils. Each pupil knowing his number can tell how he stands in relation to the other members of the class. If a new score card is posted each month, a pupil may see whether he is gaining or losing in his position in the class. If he is losing, he will be inclined to seek the reason. He may see that his neatness has a low score. This furnishes a strong incentive for work to improve in neatness. Teachers and supervisors might compare their records. The use of the card may be varied by training pupils to score their own or others' handwriting, or by one teacher calling on another teacher to score the handwriting of his pupils.

Standards. In Table XXII we give (1) standards proposed by Ayres for his "Gettysburg Edition"; (2) standards proposed by Freeman; and (3) "the Kansas Medians" which were obtained by using the directions given on page 204. Table XXII is read thus: A second-grade class should have a median score for rate of 36 letters per minute, and a median score for quality of 44, when scored by the Ayres Scale. A third-grade class should have a median quality of 47 and a median rate of 48 letters per minute. The standards for the other grades are read in the same manner.

TABLE XXII. HANDWRITING STANDARDS — RATE IN LETTERS PER MINUTE — QUALITY IN TERMS OF THE AYRES SCALE

	School grades						
	II	III	IV	V	VI	VII	VIII
Freeman standards —							
Quality	44	47	50	55	59	64	70
Rate	36	48	56	65	72	80	90
Ayres ("Gettysburg Edition") —							
Quality	38	42	46	50	54	58	62
Rate	32	44	56	64	70	76	80
Kansas medians —							
Quality	44	47	50	55	59	64	70
Rate	32	35	51	61	67	71	73

Ayres's standards represented graphically. Ayres has represented graphically his standards for the "Gettysburg Edition" as shown in Fig. 33. Quality is represented on the horizontal lines and rate on the vertical. The positions of the small circles indicate the standards for the respective grades. This plan of graphical representation is frequently helpful in interpreting the scores of a class or of a school.

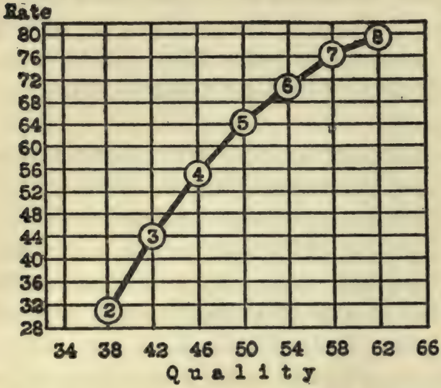


Fig. 33. Graphical Representation of Ayres's Standards for the "Gettysburg Edition" of his Handwriting Scale

The basis of satisfactory standards in handwriting. The standards of attainment are determined by two considerations: (1) they must be attainable by pupils under ordinary school conditions, and without the expenditure of an unreasonable amount of time and effort; (2) they should be high enough to assure that the pupil will have sufficient skill in writing to meet the demands which will be made upon him. These considerations are emphasized by the facts that only a limited amount of time is available for the teaching of handwriting in the ordinary school, and that after practice has progressed for a time, it does not bring as large returns as it did in its initial period.

The first of these considerations has been met by examining the handwriting of thousands of children, gathered from all parts of our country. Freeman used the results of the scoring of about five thousand specimens from each of the seven grades. These specimens were selected from a large number of specimens which were collected in fifty-six large cities of the United States. He found that the average of the scores of the upper half of these specimens gave scores for rate and quality which are approximately the standards he proposes. In checking up the second consideration, Freeman investigated the demands which are made upon those who are employed in several large commercial houses. The returns from this investigation, together with the results of the other investigation, indicated that the standards as proposed are but little more than the minimum essentials. Moreover, Freeman estimates on good evidence that these standards can be attained with an expenditure of not over seventy-five minutes a week.

Standards required for practical work. Pupils are taught to write for two reasons: (1) in order to be able to meet the practical demands for writing outside of school, and (2) in order to be able to do the writing that is required in school, particularly in high school and college. Eventually these demands will determine the standards for both rate and quality. With reference to quality Ayres[1] and Ashbaugh[2] have drawn certain conclusions from the requirements in handwriting which are set up by the examiners of the Municipal Civil Service Commission of New York City. Ashbaugh quotes a letter from the Acting Director of the commission as follows:

> I find that the Municipal Civil Service Commission of New York ordinarily uses the standard of 70 per cent as a passing grade in handwriting, but for positions where handwriting is a special requirement the standard is sometimes set at 75 per cent.

---

[1] Ayres, L. P., *A Scale for Measuring the Quality of Handwriting of Adults.* (Russell Sage Foundation, Bulletin E 138.)

[2] Ashbaugh, Ernest J., *Handwriting of Iowa School Children.* (Bulletin of the University of Iowa, March 1, 1916.)

Ayres has shown that the ratings of 70 per cent and 75 per cent, as given by the commission, correspond respectively to scores of 40 and 50 on the Ayres Scale. Since this commission recommends many persons who cannot write better than the 40 specimen of the Ayres Scale, and recommends others who write only as well as the 50 specimen, for positions where handwriting is a special requirement, it would follow that an ability to write as well as 50 on the Ayres Scale would be sufficient for all the demands which most pupils will meet.

Koos[1] has recently reported a study of the non-vocational handwriting of 1053 persons and also the handwriting of several vocational groups. He states his conclusions as follows:

> To write better than 60 is to be in a small minority (13.5 per cent of 1053 cases) as concerns handwriting ability. Moreover, four-fifths of 826 judges consider the quality 60 adequate with a generous majority approving quality 50. In the light of these facts, *it is difficult to see why, for the use under consideration (non-vocational correspondence) a pupil should be required to spend time to learn to write better than quality 60. There is even considerable justification for setting the ultimate standard at 50. As this demand touches every member of society, all children in the schools should be required to attain the standard set.*
>
> From the facts that have been presented touching the ability in handwriting of persons engaged in various occupations, it seems to the writer that *the quality 60 on the Ayres Measuring Scale for Adult Handwriting . . . is adequate for the needs of most vocations. . . . For that large group who will go into commercial work, for telegraphers, and for teachers in the elementary schools it will be necessary to insist upon the attainment of a somewhat higher quality, but hardly in excess of the quality 70.*

Standards required for school work. We have but little data on this point, but many pupils come to high schools unable to write rapidly enough for the demands placed upon them. They then often sacrifice the quality of their handwriting for the sake of greater rate. Lewis[1] examined the hand-writing of 1760 third- and fourth-year students of 166 Iowa Normal Training High Schools. He found their median score for quality to be 59.1 on the Ayres Scale, with a range from 34 to 89. Fifty per cent of the scores fell between 53.6 and 64.3. The average rate of their handwriting was 90 letters per minute. Thus, they rank with the seventh-grade standard for quality, and the eighth-grade standard for rate. Comparing their scores with those of eighth-grade children (see Table XXII), these high-school pupils write from ten to fifteen letters per minute faster, but no better than the average eighth-grade pupil. These data bear out the statement that the higher schools require greater rate of handwriting than the training of the elementary schools have furnished. Therefore, increased emphasis should be placed upon rate in teaching handwriting.

---

[1] Koos, L. V., "The Determination of Ultimate Standards of Quality in Handwriting for the Public Schools"; in *Elementary School Journal* (February, 1918), vol. 18, p. 422.

Summary. This discussion of standard scores for handwriting may be summarized by saying that there is evidence that the standards for quality given in Table XXII may be slightly higher than they should be, particularly those given by Freeman. The standards given by Ayres may be considered satisfactory. In the case of the rate of writing Freeman's standards are probably the best.

Types of situations revealed by the measurement of handwriting ability. Three types of situations which need corrective instruction may be recognized: (1) the median rate of writing is below standard; (2) the median quality is below standard; (3) the scores are too widely scattered.

Type I. Below standard in rate of writing: The Cause.

---

[1] Lewis, E. E., "The Present Standard of Handwriting in Iowa Normal Training High Schools"; in *Educational Administration and Supervision* (December, 1915), vol. 1, pp. 663–71.

Fig. 34 represents the distribution of scores for a third-grade class. The numerals along the bottom of the figure denote quality on the Ayres Scale, and the rate in terms of letters written in one minute. The numerals along the side indicate the number of pupils. A perpendicular solid line shows the location of the median for the class, and a perpendicular broken line shows the location of the standard for that grade. This class is below standard in both rate and quality. The quality will be considered under Type II. When the median rate of writing of a class is conspicuously below standard, as is the case of the third-grade class shown in Fig. 34, it is almost certain that the teacher is failing to place sufficient emphasis upon rate in his instruction. The author has found teachers and even supervisors of handwriting who admitted that they had given no attention to the rate of writing, but it was obvious that rate was important as well as quality. A few pupils are very slow in their movements, and this may account for the low rate of individual pupils, but not for a low median score except in very unusual cases.

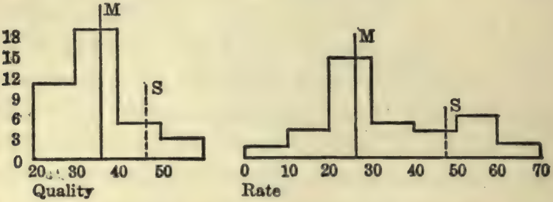


Fig. 34. Showing the Distribution of Scores in Handwriting of a Third-Grade Class.

The line M indicates the median score for the class, the line S the standard for the class.

The corrective. In considering corrective instruction for a class whose median rate score is below standard, it is necessary to bear in mind the relations which exist between rate, movement, rhythm, and quality. Investigation[1] has shown that the kind of movement, finger, arm and finger movement combined, or arm movement, does not affect the rate of writing when it is carried on for only a short time as is the case in the measurement of ability in handwriting. The apparent greater ease of production of arm or muscular movement may result in greater rate if rate is measured during a long period of writing.

Nutt has recognized what he calls "rhythm." This is a quality or characteristic of the movement. It increases with age, but has no relation to amount of arm movement or to the quality of the writing. Nutt found that rate of writing and rhythm increase together; that is, children who score high in rhythm also score high in rate, but may not use arm movement or produce a better quality of handwriting than other children.

Relation between rate and quality. Several studies have sought for a relation between rate and quality of handwriting. In the Cleveland Survey[2] it was found that "in general speed and quality vary inversely. But there is a middle series of speeds and qualities where improvement in one does not seem to interfere with the other"; that is, outside of the limits which are approximately those of the proposed standards, efforts to secure an unusual degree of quality will reduce the rate, and *vice versa*. Several investigations of adults' handwriting show that they tend to increase the rate and reduce the quality. A general view of the results bearing on this point shows that the children who write a good quality on the average write as rapidly as those who write a poorer quality. This seems to be due to the natural rhythm of the children. If this rhythm is forced or disturbed unduly the quality suffers. Thorndike's results indicate that causing a pupil to write more slowly than his normal rate did not improve the quality of the handwriting.

---

[1] Nutt, H. W., "Rhythm in Handwriting"; in *Elementary School Journal*, vol. 17, pp. 432–45.

[2] Judd, C. H., *Measuring the Work of the Public Schools*, pp. 80–81.

Drills for increasing the rate of writing. Since within limits the rate of writing may be increased without seriously disturbing the quality, it will be possible in some cases to bring the median rate up to standard by rate drills in which the pupils are caused to write at standard rates. A convenient device for doing this is represented by the following example. This is a dictation exercise arranged for the sixth grade. The rate of dictation which is indicated by the numbers printed above the words is based upon Freeman's standards. (See page 219.) The teacher should direct the class to be ready to write, then, watching the second-hand of his watch, until it is at 60, start to dictate. A little preliminary practice will make it easy to dictate the words so that they will be pronounced as indicated. For example, the teacher should be pronouncing the word "care" just before the second-hand reaches the ten-second mark, etc.

5 10 20 30
Do you take care to keep your teeth very clean, by washing
40 50 60 15
them without failing every morning and after every meal? This
20 30 40 50 60
is very necessary both to preserve your teeth a great while, and
10 20
to save you a great deal of pain. (Stop.)

At first a class will not be accustomed to this form of exercise and may not respond in a satisfactory manner, but a little patience on the part of the teacher will soon eliminate such temporary confusion. The rate of a class which is far below standard should be gradually increased. For example, if the sixth-grade class is below the fourth-grade standard, a dictation exercise arranged for the fourth grade should be used. When the class is able to respond to this satisfactorily, the dictation exercise for the fifth grade should be used and later the one for the sixth grade.

If the quality of the handwriting of certain pupils decreases because of such drills, these pupils should be excused from them or a different type of drill used. The following is a modification which will be helpful in such cases; the sentence, "The quick brown fox jumps over the lazy dog" contains thirty-five letters.

8th	-grade	pupils	should	write	this	11	times	in	4	min.	
7th	"	"	"	"	"	8	"	"	3	"	30 sec.
6th	"	"	"	"	"	6	"	"	3	"	
5th	"	"	"	"	"	5	"	"	2	"	45 "
4th	"	"	"	"	"	4	"	"	2	"	30 "
3d	"	"	"	"	"	3	"	"	2	"	10 "
2d	"	"	"	"	"	2	"	"	2	"	

The pupils should memorize the sentence and write it several times for practice and for spelling. The teacher should then time their writing. Those who do not write the required number of letters in the allotted time, as given in the table above, should be told to write faster, until they have done the test successfully.

Developing rhythm. If such exercises as described above reveal a serious sacrifice in the quality when the rate is increased, or if the pupil's handwriting cannot be brought up to standard rate, we may consider that the pupil's rhythm has not developed to the place where it will sustain this rate. Since we do not know which is the primary factor, rhythm or rate, the best procedure would be to seek to develop both. Rhythm may be increased by the use of music. If the school owns a phonograph, records suitable for use in penmanship classes may easily be secured. The time of the music may be adjusted to the grade. Careful attention to the securing of a free, well-relaxed hand position will aid in securing rate. Sometimes a careful analysis of letter forms will reveal that the pupil is forming some letters in a way that makes a satisfactory rate impossible. In such cases new forms of those letters should be taught.

Type II. Below standard in quality of writing. This condition may occur along with an unsatisfactory rate, as in Fig. 34, or when the rate is up to or above standard. In attempting to increase the median quality of the handwriting of a class, methods and devices used should be selected in the light of facts which have been established by investigations of the learning process,[1] as it occurs in learning to write. There are not sufficient data from comparative studies of different penmanship systems to establish any single system as superior to others in its effectiveness to secure results in terms of rate and quality of handwriting. Hence, the corrective to be sought is not some system of writing which is a panacea for all handwriting troubles.

General laws of learning applied. The ability to write well is a habit; hence, the laws of habit formation apply to the acquisition of this ability.

The *first* essential factor is a right start. The pupil must have a clear view of the habit to be acquired. This may mean a definite idea of the movement to be executed, or a picture of the letters or series of letters which are to be made. The start must be made with a strong initiative. Sometimes the pupil must be shocked into a desire to correct a fault of his handwriting.

The *second* essential is that of attentive repetitions. The repetitions or drills should be strongly motivated. All investigations of habit formation agree upon this point. The periods of practice are most efficient if not carried to the point of fatigue; hence, for the lower grades Freeman suggests frequent ten-minute periods of practice. In no grades should the periods be longer than twenty minutes.

---

[1] No attempt is made to review or to criticize the material which appears in numerous manuals of handwriting. Much excellent material which appears in *The Teaching of Handwriting*, by Freeman, is not even mentioned, because of lack of space. The difficulty of confining this discussion to the actual facts discovered through measurement of handwriting will be apparent.

The *third* step, as often stated, is, "Allow no exceptions to occur." If a pupil practices correct form in the penmanship class for ten minutes, and then uses poor form in a spelling class for the same length of time, the latter exercise will tend to cancel the effects of his practice in the penmanship class.

A *fourth* step is the repetition of the habit until it is well fixed. This means that the repetitions must extend beyond the point of apparent completion to permanent automatism. After this stage is reached, incentives should be found which will raise the habit from the level of mere automatism to higher levels of skill.

Motivating practice. A number of devices and plans have been proposed for the motivation of practice in correcting faults in quality of handwriting. Wilson[1] gives the result of an interesting experiment in which the Thorndike Scale was used in such a way that the students could follow their own progress in handwriting. In this case each student was competing with his own record. Several teachers have constructed scales from the specimens collected in a school or class. These scales may be constructed by rating the specimens with any one or more of the scales described. Superintendent Bliss of the Montclair, New Jersey, schools is quoted by Wilson as follows: "A scale made from the writing of pupils makes a stronger appeal than either the Ayres or Thorndike Scales." A scale, either one made from specimens collected in the school or one of those described on page 208, should be posted in the schoolroom and pupils encouraged to compare their handwriting with it frequently. For this purpose the Ayres Scale is most convenient.

---

[1] Wilson and Wilson, *The Motivation of School Work* (Houghton Mifflin Company, 1916), p. 187.

Charters [1] recommends a "writing hospital" to which the poor writers are sent until they are properly convalescent. This hospital is a special penmanship class. Stone[2] has a plan which puts all the pupils of a school in four groups for their writing lessons. These are groups 1, 2, 3, and the excused group. The special feature of this plan is that at stated intervals members of a lower group are allowed to challenge members of a higher group, and a contest for the coveted place ensues.

Many special devices for motivation are in use. Pupils write letters ordering supplies for the school, or they write invitations to school parties, pageants, etc. Some pupils write letters for the teacher or principal.

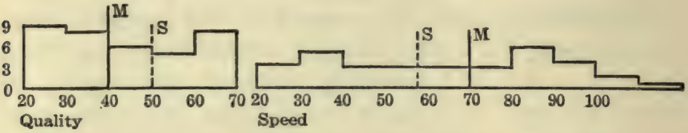


FIG. 35. SHOWING THE DISTRIBUTION OF SCORES IN HANDWRITING OF FOURTH-GRADE CLASS

The line M indicates the median score for the class, the line S the standard for the class.

Type III. Scores too widely scattered. The scores of a fourth-grade class of this type are shown in Fig. 35. As in the case of other school subjects, the pupils who are grouped together in any school grade will be found to differ widely in both rate and quality. However, these differences should be reduced to a minimum. Fig. 36 represents the scores of a fifth-grade class which exhibits what probably should be regarded as a satisfactory condition. The differences between the members of this class are much less than those of the class shown in Fig. 35.

---

[1] Charters, W. W., *Teaching the Common Branches.* (Houghton Mifflin Company, 1916.)

[2] Stone, C. R., "Motivation of the Formal Writing Lesson Through a Special Classification of Pupils for Writing"; in *School and Home Education,* June, 1915.

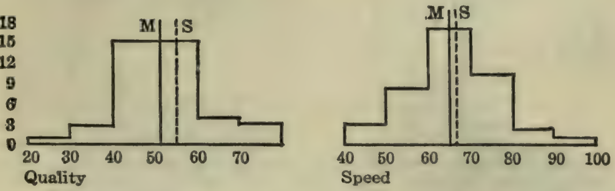


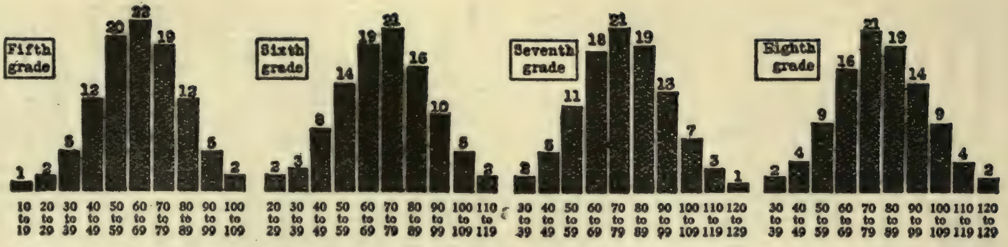
FIG. 36. SHOWING THE DISTRIBUTION OF SCORES IN HANDWRITING OF FIFTH-GRADE CLASS

The line M indicates the median score for the class, the line S the standard for the class.

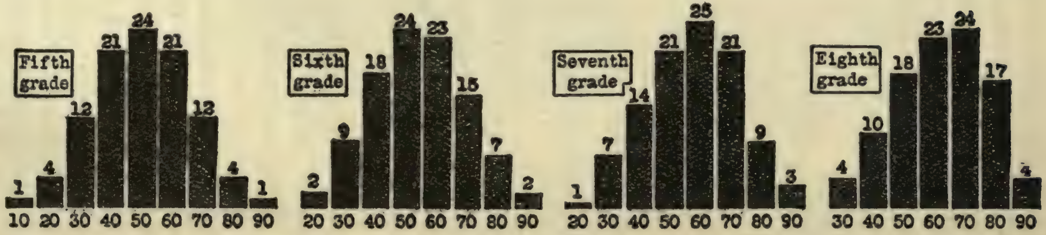
In connection with his "Gettysburg Edition," Ayres has given standard distributions for the four upper grades. These are reproduced in Fig. 37. The teacher may use them to ascertain whether or not the scores of his class are too widely scattered.

Correctives. The reduction of a high degree of individual differences is largely a matter of dealing with individual pupils. A reclassification may be wise where it is possible, but for the most part the classification of pupils is determined by their standing in other subjects. Those pupils who are distinctly above the eighth-grade standard should be excused from the penmanship class. They may spend the time thus saved upon other subjects. Dictation exercises, such as described on page 226, will tend to reduce the degree of individual differences in rate of writing.

Diagnostic measurements. In the case of pupils who are below standard in quality, it is helpful to diagnose their



Rate of writing. Columns represent per cent of pupils in each of four upper grades commonly found to write the number of letters per minute shown below each column.



Quality of writing. Columns represent per cent of pupils in each of four upper grades commonly found to have handwriting of qualities shown below each column.

FIG. 37. SHOWING STANDARD DISTRIBUTION OF SCORES IN HANDWRITING. (*After Ayres.*)

handwriting using either Freeman's Scale or Gray's Score Card. This will give the teacher a statement of the particular defects which exist and this information will provide a basis for prescribing corrective instruction. As typical of this procedure we quote the following:[1]

> A detailed analysis of the faults which appear in the child's writing and of the adjustments which are necessary to correct them has been worked out by Mr. C. W. Reavis, Principal of the Laclede School, St. Louis, Missouri, on the basis of his experience in supervision, and is here presented with his permission.

Analysis of Defects in Writing and their Causes, in use by Principal Reavis

Defect	*Cause*
Too much slant	1. Writing arm too near body. 2. Thumb too stiff. 3. Point of nib too far from fingers. 4. Paper in wrong position. 5. Stroke in wrong direction.
Writing too straight	1. Arm too far from body. 2. Fingers too near nib. 3. Index finger alone guiding pen. 4. Incorrect position of paper.
Writing too heavy	1. Index finger pressing too heavily. 2. Using wrong pen. 3. Penholder of too small diameter.
Writing too light	1. Pen held too obliquely or too straight. 2. Eyelet of pen turned to side. 3. Penholder of too large diameter.
Writing too angular	1. Thumb too stiff. 2. Penholder too lightly held. 3. Movement too slow.

---

[1] Freeman, F. N., *The Teaching of Handwriting* (Houghton Mifflin Company), pp. 71–72.

Writing too irregular	1. Lack of freedom of movement.
	2. Movements of hand too slow.
	3. Pen gripping.
	4. Incorrect or uncomfortable position.
Spacing too wide	1. Pen progresses too fast to right.
	2. Too much lateral movement.

QUESTIONS AND TOPICS FOR STUDY

1. A teacher may judge the handwriting of his class by watching the pupils while they write or by examining the specimens which they have written. Which is the better method if the purpose is to make comparisons of classes? Which is better for discovering the handwriting defects of individual pupils? What factors would you keep in mind in watching children while they write? What factors in the other method?
2. Ask a class to write the three sentences from Lincoln's *Gettysburg Address*. Direct them to start together and write as *rapidly* as they can for one minute. At the end of one minute stop them and direct them to record the number of *letters* they have written. Then ask them to begin again and write for one minute writing as *well* as they can. If you wish to eliminate practice effects, repeat the experiment, again reversing the order of the directions. Note the difference in the rates due to the nature of the directions.
3. Select ten or preferably one hundred specimens of handwriting and rate them every day for several days by means of the scale you have. Keep the record of your day's rating, but do not use them to help you in making future ratings. After several ratings note the consistency of your ratings.
4. Use the Gray Score Card (or Freeman Scale) in scoring the poorer specimens of handwriting. Prescribe the drills you would use in correcting these defects. Compare this with the recommendations of other teachers or students. Try your prescription on the pupils concerned if possible.
5. For what purpose would you use the dictation exercises?
6. Select a defect of letter formation frequently found in a pupil's handwriting. Direct the pupil's attention to this defect and challenge him to correct it. Direct that a record be taken as follows: If the defect were found in letter "a" instruct the pupil to count the number of such errors to be found in fifty consecutive "a's" as they occur in his handwriting written prior to the time you pointed out the defect. After a period of practice, direct the pupil to make another counting from his handwriting written at some period other than the writing period.

CHAPTER IX

THE MEASUREMENT OF ABILITY IN LANGUAGE AND GRAMMAR

The measurement of ability to write compositions. The plan for measuring the ability of pupils to write compositions is very similar to that used for handwriting. Composition scales have been devised which consist of a number of compositions arranged in order of merit, and a pupil's composition is measured by "matching" it with the composition of the scale which most nearly equals it in merit. As in the case of handwriting, care must be exercised in securing compositions from pupils. The first draft of a composition is frequently inferior to the form obtained when it has been rewritten. Also there will probably be a difference between compositions written as a class exercise and compositions prepared as home work.

The Willing Composition Scale for measuring compositions written as a class exercise. Willing[1] has devised a scale which consists of compositions written as a class exercise. The topic was "An Exciting Experience." Several particularly exciting experiences were suggested by the teacher and the pupils were allowed twenty minutes for writing. The compositions were rated both for form (errors in spelling, punctuation, capitalization, and grammar) and for "story value." Those chosen for the scale increase gradually in both form and "story value." The scale is reproduced here so that teachers may understand better this type of measuring instrument in the field of composition.

---

[1] Willing, M. H., "The Measurement of Written Composition in Grades IV to VIII"; *English Journal* (March, 1918), vol. 7, p. 193.

WILLING SCALE FOR MEASURING WRITTEN COMPOSITION

(The values: 90, 80, 70, 60, 50, 40, 30, and 20 given the respective samples are arbitrary and merely for practical convenience. 20 means 15 to 24.9, 30 means 25 to 34.9, etc.)

20

Deron the summer I got kicked and sprain my arm. And I was in bed of wheeks And it happing up to Washtion Park I was going to catch some fish. And I was so happy when I got the banged of I will nevery try that stunt againg

Number of mistakes in spelling, punctuation, and syntax per hundred words, 30.

30

The other day when I was rideing on our horse the engion was comeing and he got frightened so he through me down and I broke my hand.

And the next thing I done was I went to the doctor and he put some bandage on it and told me to come the next day so I came the next day and he toke the bandage off and he look at it and then it was better.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 23.

40

My antie had her barn trown down last week and had all her chickens killed from the storm. Whitch happened at twelve oclock at night. She had 30 chickens and one horse the horse was saved he ran over to our house and claped on the dor whit his feet. When we saw him my father took him in the barn where he slepped the night with our horse. When our antie told us about the accident we were very sorry the next night all my anties things were frozen. The storm blew terrible the next morning and I could not go to school so I had to stay home the whole week.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 17.

50

One time mother and father were going to take sister and I for a long ride thanksgiving, We had to go 60 miles to get there, When sister and I herd about it we were very glad. It was a very cold trip. We four all went in a one seated automobile. Dady drove and mother held me and sister sat on the top the top was down. Mother could not hold sister for she was two heavy. When we got there they had a hot fire ready for us and a goose dinner. We were there over night. In the morning it was hot out. This was on a farm. Sister and I got to go horse-back riding. It was lots of funs. They had children. The children were very nice. Our trip home was very cold. When we got home it had snod.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 14.

60

One time when mother, some girl friends and myself were staying up in the mountains. An awful storm came up. At the we were way up the mountain. The lightning flashed and the thunder roared. We were very frightened for the cabin we were staying at was at the foot of the mountain. We did n't have our coats with us for it was very warm when we started. There were a few pine trees near us so we ran under them. They did n't do much good for the rain came down in torrents. The rain came down so hard that it uprooted one of the trees. Finely it began to slack a little, So we thought we would try and go back. About half way down the mountain was a little hut. We started and when got about half way down it began to rain all the harder. We did n't know what to do for this time there was n't any trees to get under. We decided to go on for the nearest shelter was the hut. Finely we got there cold and wet to the skin.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 11.

70

When I was in Michegan I had an exciting thing happen or rather saw it, it was when the big steamship plying between Chicago and Muskegon was sunk about 7 o'clock in the evening. It caught on fire with a load of cattle and products from the market on board, one of the lifeboats carrying some of the people who were on board landed at our pier. The "Whaleback" steamer which goes between Chicago and Muskegon was two hours later in coming than the freighter and was stopped to clear up the wreckage. all of the cattle and products and an immense cargo of coal were lost, but there were only two people lost. the ship tried hard to get to port with her cargoe but, could not reach it. The next morning we found planks, and parts of the wreck on the beach. Our cottage was at the top of a cliff and it was just one hundred feet to the lake from our cottage, we had a beautiful view, and the sight of the fire on the horizon was a beautiful sight (though it was pitiful).

Number of mistakes in spelling, punctuation, and syntax per hundred words, 8.

80

Near our ranch in Fort Logan there was a chicken ranch. One day my sister and I went up to the chicken ranch on our horses. Coming back there was a road leading from our house to the main road and along this road were half rotted stumps. On every one of these stumps what do you think we saw. We saw snakes! snakes! snakes! I suppose these snakes were shedding their skins, they were of every color, shape, and size. But when sister and I saw these snakes we whipped our horses into a gallop and away we went just as hard as we could go. When we got to the house we went in and mamma could n't get us out of the house that day. I was so scared that I believe I dreamed about snakes for a month.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 5.

90

The most exciting experience of my life happened when I was but five years of age. I was riding my tricycle on the top of our high terrace. Beside the curbing below, stood a vegetable wagon and a horse. Suddenly I got too near the top of the terrace. The front wheel of my tricycle slipped over and down I went, licety-split, under the horse standing by the curbing. I had quite a high tricycle and the handlebars scraped the horse's stomach, making him kick and plung in a very alarming manner. I was directly under him during this, but finally rolled over out of his way and scrambled up. I looked at my hands! Most of the first finger and part of the thumb of my left hand were missing. The horse had stepped on them. I had endured no sensation of pain before this, but now my mangled hand began to hurt terribly. I was hurried to the hospital and operated on, and now you would hardly notice one of my fingers is missing. I certainly have good cause to congratulate myself on my good fortune in escaping with as little injury to myself as I did, for I might have been terribly mangled in my head or body.

Number of mistakes in spelling, punctuation, and syntax per hundred words, 0.

DIRECTIONS FOR USING WILLING'S SCALE FOR WRITTEN COMPOSITION

In using the Composition Scale, these directions should be followed carefully because the compositions were written by school children who followed these same directions.

1. The teacher should make certain that all pupils are provided with good pen points and ink, or well-sharpened pencils if pencils are to be used. Have distributed to each pupil two sheets of theme paper (approximately 8½ by 11). It is best to use theme paper which has printed at the top the suggested list of topics. If this kind of paper is not used, the teacher must write the following list of topics on the blackboard:

An exciting experience.
A storm.
An accident.
An errand at night.
A wonderful story.
An unexpected meeting.
In the woods.
In the mountains.
On the ice.
On the water.
A runaway.

2. The teacher should then say to the pupils: "I want you to write me a story. It is to be a story about some exciting experience that you have had, about something or other very interesting that has happened to you. If nothing of the sort has ever happened to you, then tell me of an exciting experience some one you know has had. You may even make up a story of this kind, if you have to, though I believe you will do better, on the whole, with a real one. I am going to give you about twenty minutes in which to write. You are to write on both sides of the paper, to do all the work yourselves, and to ask no questions at all after you begin. You may make whatever corrections you wish between the lines. There will be no time to rewrite your story.

"The general subject together with some suggestions is printed at the top of the paper on which you are to write. (I have written the general subject on the blackboard, together with some suggestions.) You do not have to write on any of these topics unless you want to; they are merely to help out in case you cannot think of an exciting experience yourself. You may begin now as soon as you wish."

3. Allow opportunity for asking questions and make an effort to put the children at ease. Allow full twenty minutes for the actual writing. At the end of this time say to the pupils:

"You are to have four or five minutes in which to finish your stories; make corrections and count the number of words written. Write this number at the end of your story, write also your name, age, and grade."

At the end of five minutes collect the papers.

It is important that the pupils be not allowed to correct their compositions, except such corrections as they may make during this period of four or five minutes. The teacher must remember that it is this type of composition which was used in making the scale and establishing the standards.

4. In rating the compositions by means of the scale, two qualities are recognized: (1) "story value" and (2) "form value." The composition should be rated for "story value" first.

Rating for "Story Value." Read the compositions, neglecting all errors of grammar, punctuation, capitalization, and spelling, and keeping in mind only the value of the story which the pupil is telling. As the compositions are read, sort them into piles, placing in one pile those which most nearly resemble in "story value" composition 20 of the scale; in another pile those which most nearly resemble composition 30; in another pile those which most nearly resemble composition 40; and so on for the other compositions of the scale.

After this is done, compare the compositions in each pile with each other, in order to make sure that the rating has been done correctly. Make any adjustments which you think should be made. Mark each composition with the value of the scale composition which it most nearly resembles. In case you believe that the true "story value" of the pupil's composition lies between that of two of the scale compositions, the interpolated marks 25, 35, etc., may be used.

Rating for "Form Value." After the compositions have been rated for "story value," carefully mark all errors in grammar, punctuation, capitalization, and spelling. Count these errors and multiply the total by 100. Divide this number by the number of words in the composition. This quotient is the number of errors per hundred words.

The quotient found, as directed above, and the "story value" of the composition constitute the pupil's score. These scores are valuable to the teacher. They show the standing of each pupil in two respects. They tell the teacher whether the pupil needs to give attention to the "form" of his writing (grammar, punctuation, capitalization, and spelling) or to the "story value," or to both.

Recording the scores. For recording the scores the record sheet shown in Fig. 38 is used. Sort the compositions for a class into piles according to their "story value." (By a class we mean the pupils who belong to the same grade and who recite together. If a teacher has a class composed of pupils belonging to two grades — say some belonging to 5 A and some belonging to 6 B — it will be necessary to make two tabulations.)

If interpolated values have been used, they should be grouped according to the explanation of the scale value which is given at the top of the scale.

Take the compositions whose "story value" is 20. These are to be listed in the first column of the table in the space which corresponds to their "form value." For example, in the first space of this column record the number of these compositions whose "form value" is between 0 and 2.9; in the second space record the number of compositions whose form value is between 3 and

Class Record Sheet

City ....School.........................Grade.............

Teacher ....Date.......................191....

Errors per Hundred Words	Story value								Distribution for Errors
	20	30	40	50	60	70	80	90	
0 to 2.9.......................	....	....	....	....	....	....	....	....	....
3 to 5.9.......................	....	....	....	....	....	....	....	....	....
6 to 8.9.......................	....	....	....	....	....	....	....	....	....
9 to 11.9.......................	....	....	....	....	....	....	....	....	....
12 to 14.9.......................	....	....	....	....	....	....	....	....	....
15 to 17.9.......................	....	....	....	....	....	....	....	....	....
18 to 20.9.......................	....	....	....	....	....	....	....	....	....
21 to 23.9.......................	....	....	....	....	....	....	....	....	....
24 to 26.9.......................	....	....	....	....	....	....	....	....	....
27 to 29.9.......................	....	....	....	....	....	....	....	....	....
Above 30.......................	....	....	....	....	....	....	....	....	....
Distribution for "story value"..	....	....	....	....	....	....	....	....	....

Class Medians. Form value............ Story value............

FIG. 38. SHOWING CLASS RECORD SHEET FOR USE WITH WILLING'S COMPOSITION SCALE

5.9; in the third space record the number of compositions whose "form value" is between 6 and 8.9, etc. After the compositions in each pile have been recorded in this way, the number of compositions recorded on each line should be counted and the total entered in the total column. The same should be done for the compositions entered in each column.

Finding the class scores. The median score for form value may be found by arranging the compositions in order of the form scores and taking the score of the middle composition. In case there is an even number of papers, the average of the scores on the two middle ones should be taken. The median score for "story value" may be found in the same way. The median scores may also be calculated from the distributions by following the directions given on page 103.

Tentative standards. Tentative standards for Willing's Composition Scale are given in Table XXIII. It will be noticed that the median scores for Denver are conspicuously below those for five Kansas cities. This may be due to the fact that reports have been received from only a few cities.

TABLE XXIII. MEDIAN SCORES FOR WILLING'S COMPOSITION SCALE

Grade	*Denver*		*Five Kansas cities*	
	Story value	*Form value*	*Story value*	*Form value*
IV	32	22	44	12
V	43	16	58	10
VI	50	14	75	5
VII	60	11	77	5
VIII	63	10	82	6

Other scales for measuring written composition. A scale called the *Nassau County Supplement* has been devised by Trabue. It consists of nine compositions, seven of which were written by elementary-school pupils on the topic, "What I should like to do next Saturday." It is designed to measure only "story value" of compositions. Copies of this scale may be obtained from the Bureau of Publications, Teachers College, Columbia University, New York City.

The Hillegas Scale consists of ten compositions ranging from an artificial production whose scale value is zero to the tenth composition whose scale value is 9.3. Three of the ten compositions are artificial productions, five were written by high-school pupils, and the remaining two by college freshmen. No two were written on the same topic and they vary greatly in length and type. In the *Thorndike Extension of the Hillegas Scale*, only a few of the compositions of the original scale have been used and several compositions are given for each degree of merit in the middle of the scale. Twenty-nine compositions represent fifteen degrees of merit within approximately the same range as the original scale. This makes a more finely divided scale than the original one. Copies may be obtained from the Bureau of Publications, Teachers College, Columbia University, New York City.

The Harvard-Newton Composition Scale consists of four separate scales, one for each form of discourse; argumentation, description, exposition, and narration. Each of the scales consists of six compositions written by eighth-grade pupils and arranged in order of merit as determined by the marks assigned by teachers rating them as eighth-grade compositions. For each composition there is given a statement of the most significant merits and defects. Copies of the scale may be secured from the Harvard University Press, Cambridge, Massachusetts.

The compositions used by Breed and Frostic[1] in deriving their scale were written by sixth-grade pupils under uniform conditions. A part of the story called "The Picnic" was read to the class and they were given twenty minutes to complete it. The method of selecting compositions for the scale and determining scale values was similar to that employed by Hillegas.

---

[1] Breed, F. S., and Frostic, F. W., "A Scale for Measuring the General Merit of English Composition"; in *Elementary School Journal*, vol. 17, pp. 307–25.

The measurement of ability in English Grammar. Charters's Diagnostic Test in Language and Grammar for Pronouns. Charters collected more than twenty-five thousand errors that pupils make in using pronouns in their oral language. These were classified under forty heads; that is, there were only forty different kinds of errors in the use of pronouns in the total twenty-five thousand. The language part of the test consists of eighteen sentences. The pupils are required to write the correct form. This test is designed to be used in grades three to eight. In the grammar part of the test, which consists of twenty-four sentences, they are required to give the reason for making the correction. The form of this test is illustrated by a few of the exercises given below. The amount of credit to be given for doing each exercise correctly has been determined. This test measures two abilities: (1) The ability to use correct forms of pronouns. The measure of this ability is his "language score" and is the sum of the values of the language exercises done correctly. (2) The ability to give the grammatical rule which tells which form is correct. The measure of this ability is his "grammar score" and is the sum of the values of the grammar exercises done correctly.

Recording the scores. For this purpose a class record sheet with detailed directions is furnished with the tests.

Recording scores for purpose of diagnosis. In order to obtain a diagnosis of the abilities of the pupils of a class, the form of tabulation which is partly shown in Fig. 38 is helpful. It gives in a compact form the record of each pupil

	Language value of exercise		*Grammar value of exercise*
8. Who do you want ?.....................	2	8.	2
....		....	
16. They made baskets and filled it with holly...	1	16.	3
....		....	
28. Was that him?.....................	3	28.	3
....		....	

SAMPLE EXERCISES FROM CHARTERS'S DIAGNOSTIC TEST IN LANGUAGE AND GRAMMAR FOR PRONOUNS

Name of pupil	*Number of Exercises*																		
	2		3		4		5			39		40		41		42		*Total*	
	L	*G*	*L*	*G*	*L*	*G*	*L*	*G*		*L*	*G*	*L*	*G*	*L*	*G*	*L*	*G*	*L*	*G*
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
Total........	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
Per cent correct....	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...
Standard...	...	...	...	...	...	...	...	...	......	...	...	...	...	...	...	...	...	...	...

FIG. 39. RECORD SHEET FOR DIAGNOSIS

Charters's Diagnostic Test in Language and Grammar. Exercises 1 and 29 are omitted because they are correct sentences.

on each exercise. With this tabulation before him the teacher can determine (1) what errors should be given more emphasis and (2) what pupils are lacking in ability. Since the test includes all the pronoun errors, the teacher may be sure that his pupils have been tested completely in this field.

Standards. No standards are yet available for this test, but those desiring to use it may obtain the standards from the Bureau of Educational Research, University of Illinois, Urbana, Illinois, as soon as they have been determined.

Starch's Punctuation Scale. Starch has devised a punctuation scale, which consists of a number of sentences which the pupil is to punctuate correctly. The sentences are grouped in exercises of gradually increasing difficulty of punctuation. The nature of the scale may be illustrated by the following extracts. The pupil's score is the value of the highest step which he does seventy-five per cent correct.

Step 6

1. We visited New York the largest city in America.
2. Everything being ready the guard blew his horn.
3. There were blue green and red flags.
4. If you come bring my book.

Step 7

1. I told him but he would not listen.
2. Concerning the election there is one fact of much importance.
3. The guests having departed we closed the door.
4. The train moved swiftly but Turner arrived too late.

Step 10

1. A tall square building is located on State Street.
2. Washington Irving whose personality was genial and charming became very popular in England.
3. You see John how I stand.
4. On the path leading to the cellar steps were heard.

Step 13

1. I saw no reason for moving therefore I stayed still.
2. There are three causes poverty, injustice and indolence.

Step 16

1. As in warfare a band of men though strong and brave individually is collectively weak if it is not well organized so a speech a report an editorial an essay any composition though its parts may be forcible or clever is weak as a whole if it is not well organized.

Standards. The following are tentative standards of attainment for the ends of the respective school years:

Grade	*Score*
Seventh	8.0
Eighth	8.3

No diagnosis obtained. Starch's Punctuation Scale does not yield a diagnosis because he did not analyze the field of punctuation to determine the types of sentences requiring punctuation. In this respect, as well as in others, it differs from Charters's Diagnostic Test in Language and Grammar described above.

Measuring accuracy in copying. Copying is a phase of school work which receives little explicit attention. This is probably due to the assumption that pupils are able to copy accurately because it appears to be such a simple activity. Copying bears a relation to written expression and to other school subjects as well. Themes are usually copied before being submitted to the teacher. In solving problems in arithmetic the quantities are copied from the text. In gathering information from references copying occurs.

The Boston test. The following test of pupils' ability to copy printed matter was prepared by a group of Boston[1] teachers:

---

[1] *Determining a Standard in Accurate Copying.* (Boston Public Schools, English, School Document no. 2, 1916.)

Directions for Giving and Scoring the Test

1. Read to the pupils the directions which are printed at the head of the selection they are to copy, but give them no further help. For example, do not specify possible errors which may be made.
2. Pupils ought not to see the selection until they are ready to copy it. Hence it should be placed on the desk face down until the signal is given to begin work.
3. Every error should be checked distinctly.
4. The errors which were to be noted were as follows: In spelling, capitalization, punctuation, undotted "i's," uncrossed "t's"; in omitting words, in adding words, in wrong words used, and in misplaced words.

Directions to Pupils

Copy in ink as much of the following selection as you can copy accurately in fifteen minutes without hurrying. Accuracy is more important than speed:

Lieutenant Ouless

In this story a young British lieutenant, in a moment of extreme irritation, strikes a private soldier. The act is one that calls for dismissal from the Queen's service. What is the officer to do? He cannot send money to the soldier — who happens to be the redoubtable Ortheris himself — nor can he apologize to him in private. Neither can he let matters drift. Ortheris, too, has his own code of pride and honor; he too is a "servant of the Queen"; but how is the insult to be atoned for? The way out of this apparently hopeless muddle is a beautifully simple one, after all. The lieutenant invites Ortheris to go shooting with him, and when they are alone, asks him "to take off his coat." "Thank you, sir!" says Ortheris. The two men fight until Ortheris owns that he is beaten. Then the lieutenant apologizes for the original blow, and the officer and private walk back to camp devoted friends. That fight is the moral salvation of Lieutenant Ouless.[1]

Kinds of errors made. This test was given to 4494 first-year pupils in the Boston High Schools in November, 1914, and therefore may be considered to measure the ability of pupils completing the eighth grade. The results are both interesting and significant. The following is quoted from the Bulletin mentioned above:

---

[1] Bliss Perry, *A Study of Prose Fiction.*

> The errors noted consisted of nine different kinds, and the number of each kind made in this test by 4494 pupils is shown by the following tabulation:
>
> | | |
> |---|---|
> | Spelling | 5,829 |
> | Capitalization | 644 |
> | Omitted words | 4,077 |
> | Added words | 606 |
> | Wrong words used | 840 |
> | Misplaced words | 105 |
> | Punctuation | 5,876 |
> | Undotted "i's" | 8,794 |
> | Uncrossed "t's" | 606 |
> | Total | 27,377 |
> | Average errors per pupil | 5.54 |
>
> *Misspelled words.* The test consisted of 170 words, 105 of them different words. It is a notable fact that every word was misspelled by somebody. It is also interesting that 92.2 per cent of the words in the test are found in Jones's *Concrete Investigation of the Material of English Spelling.*[1] In spite of the fact that these are words commonly used by children in their writing, 11.8 per cent of them were misspelled more than 100 times. This does not mean that 11.8 per cent of the children missed these words, because one pupil might have missed the same word more than once.
>
> It is impossible to make any statement in regard to the average because many of the words occur in the selection more than once, and if misspelled by the same person each time it occurs it is counted more than one error. Some children spelled a word incorrectly in one place and correctly in another. One boy spelled "lieutenant" wrong four out of five times, and spelled it a different way each time. Then, not all the children finished the entire selection, and no record was kept of the exact number of words each

---

[1] Published by the University of South Dakota.

wrote. However, 4494 pupils taking the test made 5829 errors in spelling alone, the number of errors for each word varying from 1 to 1045.

Undotted "i's" and uncrossed "t's." The errors made by leaving the "i's" undotted and the "t's" uncrossed comprise about one third of the entire number of errors and are largely important because of their value to legibility, as pointed out by Ayres. In connection with these errors, it is very noticeable that most of them were confined to comparatively few pupils. If a child showed a tendency to dot his "i's" and cross his "t's" in the first few lines, the chances were that that individual would have but few errors. On the other hand, if the child made many errors in the first part of the paper, there were many throughout the copying. One boy went through the entire paper without dotting an "i." Many others dotted only a small part of them.

The same test was given in Kansas City, Missouri, to the pupils in the seventh grade and in the first year of the high school. (Kansas City has only seven grades below the high school.) The average errors per pupil were 8.04 in the seventh grade, and 6.83 in the first year of high school.

Remedying the situation revealed. When a teacher learns the specific language weaknesses of his pupils, he is then in position to apply more intelligently his stock of methods and devices of instruction. In language, as in the case of the other subjects, the teacher must instruct individual pupils who are grouped together rather than groups of pupils. Furthermore, each pupil should receive the instruction which he needs to correct his language errors.

If pupils are weak in a language ability, such as punctuation, the laws of habit formation apply. After being sure that he understands the function of the punctuation marks, a pupil must have practice in punctuating his own writing. This probably is not sufficient. Exercises for practice can be constructed by taking appropriate material and reproducing it without the punctuation marks.

Until a teacher recognizes definite and specific ends to be attained, there is certain to be a large degree of dissipation of his efforts. Perhaps one reason why language instruction so often does not produce satisfactory results is that it is not directed toward the engendering of definite abilities. That our present standards of language are chaotic is indicated in the report of a recent investigation.[1]

Present standards for composition indefinite. Six compositions were typewritten without any identifying marks. They were "graded" on the scale of 100 per cent by twenty-four eighth-grade teachers who were asked to follow certain typewritten directions. The six compositions were then "completely corrected so far as mechanical or measurable errors were concerned." The corrected compositions were graded by the same teachers according to the same directions.

If the "mechanical errors" of the compositions were significant factors in determining the first set of marks, the second set of marks should be conspicuously higher. However, this was not the case. For two of the compositions the average "grade" was less after the "mechanical errors" had been corrected. The individual marks show that some teachers consider form important, and that others tend to disregard it in marking a composition.

Keeping a record of pupils' errors. In teaching spelling, teachers have kept a record of pupils' errors and have emphasized these words in their teaching. In our consideration of spelling it was urged that teachers first ascertain what words their pupils were unable to spell correctly. This plan may be adapted to the teaching of other aspects of language. The teacher should ascertain the pupils' grammatical errors, and then equip them with the rules of grammar which are needed to correct them. This has been done on a large scale in St. Louis and Kansas City, Missouri.[1]

---

[1] Brownell, Baker, "A Test of the Ballou Scale of English Composition"; in *School and Society*, vol. 4, pp. 938–42.

The point of view in locating errors and applying correctives is most important. Perhaps the scales and tests described in this chapter will have fulfilled their most important function if they cause teachers to analyze and define "language ability" in more specific terms. It is believed that their use will tend to produce this result, especially such a test as Charters's Diagnostic Test in Language and Grammar for Pronouns which is based upon an analysis of that field. Analysis of "language ability" and specific definition of the elements are greatly needed. Upon the accomplishment of these two things depends the construction of more valuable measuring instruments in the language field and the scientific determination of methods and devices of instruction.

QUESTIONS AND TOPICS FOR STUDY

1. Give the copying test to your pupils following the directions carefully. Do the results agree with your estimate of the ability of your pupils to copy?
2. Keep accurate lists of the language errors of your pupils both oral and written. What are the rules which are necessary to correct these errors? Are they the rules upon which you are placing the most emphasis in your teaching?
3. Do you have definite objective standards of attainment in English composition? Can you use the scales described in this chapter to establish such standards?
4. Do you think pupils would be helped by having definite objective standards of attainment established for them?
5. Secure a copy of Willing's Composition Scale and post it in the classroom. Have pupils measure their compositions with it.
6. Why is Charters's Diagnostic Test in Language and Grammar more helpful to the teacher than Starch's Punctuation Scale?
7. What do you think of Charters's method of determining what exercises to use? Is it a good method? Why?

---

[1] See report by W. W. Charters in the *Sixteenth Yearbook of the National Society for the Study of Education*, part I.

CHAPTER X

THE MEASUREMENT OF ABILITY IN GEOGRAPHY AND HISTORY

GEOGRAPHY and history are different from the school subjects treated in the preceding chapters. Subjects such as reading, the operations of arithmetic, spelling, and the like are sometimes called "tool subjects" to distinguish them from such subjects as geography and history which are called "content subjects." Silent reading is a tool which a pupil uses in studying geography and history. The operations of arithmetic are tools which are used in solving problems.

Several tests have been devised for both geography and history, and, although they are open to criticism, they are more effective as a measuring instrument than tests or examinations prepared by the teacher. The questions have been very carefully selected, have been evaluated, and the tests have been standardized. In both geography and history there are a very large number of items of information. Some of these are important, while others are unimportant. Authorities agree on the importance of some facts; on others they disagree. Hence, the selection of the questions is very important. On page 10 we found that questions were not equally difficult, and hence it is important to have the amount of credit to be given for each question scientifically determined. Finally, pupils' scores cannot be interpreted without standards.

The criticism is frequently made, that while it is possible to measure "what a pupil remembers," it is not possible to measure his ability to answer "thought questions." This statement is not true. It is possible to measure the ability of pupils to think. However, it is very significant that investigation has shown that there is a very definite connection between a pupil's ability to remember and his ability to think. One investigator[1] in history showed that there was a very close agreement between a pupil's score on a memory test and his score on a thought test. This result is just what we should expect when we recall that reasoning involves the use of facts and a person cannot reason effectively unless he has command of the necessary facts. The application of this relation between ability to remember and ability to think is that in using a memory test, we are also indirectly measuring the ability of pupils to think in the same field. For the most part tests in geography and history have been devised so recently that we do not have proof of their value as we do in reading and arithmetic. However, certain ones of the available tests give promise of being helpful to the teacher. In this chapter we will describe two tests in geography and one test in American history.

I. Geography

1. Courtis's Standard Tests in Geography for States and important cities of the United States. This test is explicitly designed to cover only two topics of geography, but all teachers will probably agree that they are important ones. The plan of the test is to provide each pupil with an outline map of the United States showing the boundaries of the several States. Each State is given a number. The first part of the test consists of answering for each State the question, "On the map above what is the number of ....?" In the second part of the test the pupil is asked to give the number of the State in which certain cities are located. The preliminary test which is given, so that the pupil may understand just what he is to do, is reproduced in Fig. 40, to illustrate this type of test.

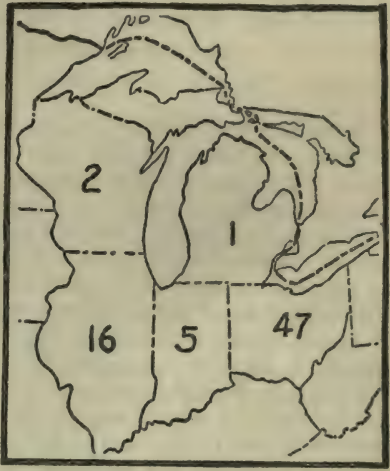
---

[1] Buckingham, B. R., "Correlation between Ability to Think and Ability to Remember, with Special Reference to United States History"; in *School and Society* (April 14, 1917), vol. 5, p. 443.

INSTRUCTIONS

Write after each state the number printed in that state, on the map at the right. For instance: Write 1 after Michigan. What number should be written after Ohio?

In the same way, after each city write the number of the state in which that city is located. Write 1 after Detroit. What should be written after Chicago?



STATE			**CITY**		
Questions	**Number**		**Questions**	**Number**	
1. Michigan ?.......	....	....	**1. Detroit ?**.........	....	....
2. Ohio ?............	....	....	**2. Chicago ?**........	....	....
3. Indiana ?........	....	....	**3. Cleveland ?**......	....	....
4. Illinois ?.........	....	....	**4. Indianapolis ?**...	....	....
5. Wisconsin ?......	....	....	**5. Madison ?**........	....	....

FIG. 40. ILLUSTRATING COURTIS'S STANDARD TEST IN GEOGRAPHY FOR STATES AND IMPORTANT CITIES OF THE UNITED STATES

Marking the test papers and recording the scores. Detailed directions for marking the test papers and for recording the scores are furnished with the test and hence need not be given here.

2. Hahn-Lackey Geography Scale. A different type of test has been devised by H. H. Hahn and E. H. Lackey. This scale consists of questions which were very carefully selected. The plan of selection is described as follows :[1]

> Since texts will be used by a large majority of teachers for years to come, our primary purpose was to construct a scale for the testing of the teaching of geography from *text*-books. But when we realized that not one but a number of texts are being taught, we had to modify our plan. Our first modification consisted of limiting our questions to the phases of geography treated in common by six modern texts. Then we found that some of these phases were treated more fully by some authors than they were by others. A second modification of our plan was, therefore, necessary; namely, to select the common subject-matter, or, in other words, the esentials of subject-matter in each phase. In the selection of the essentials of subject-matter, the common subject-matter in these texts was largely our guide, but we also checked our exercises by principles and minimum essentials as they have been worked out by makers of geography curricula. (See 1914 and 1916 *Yearbook* [of the National Society for the Study of Education].) Over six hundred questions and exercises were selected by three teachers, covering this common subject-matter. These exercises were then examined by the authors of the scale, first, with reference to repetitions, and duplications were eliminated. They were next examined for language difficulty. The wording of many of the exercises was changed, some of them were actually tried out on children, and in many instances technical expressions which would convey exact meaning to mature students of geography were eliminated and the ordinary language of children substituted. This is particularly true of the exercises in the lower reaches of the scale. The exercises intended for the upper reaches of the scale were not freed from technical expressions the meaning of which pupils are

---

[1] From an unpublished account of the derivation of the scale by the authors.

expected to know as evidence of geography ability. Thus we find such expressions in the scale as "the Fall Line," "climate," "continent," "natural wonders," "natural geographic barriers," "agencies," "cyclonic storms," and many others equally as technical. The exercises were examined, in the third place, as to their scope, as suggested before. Nothing was included beyond the essentials of geography. Finally, the list of exercises was revised so that it contained about an equal number of memory and thought questions.

These questions were classified according to difficulty by giving them to 1696 pupils in twelve schools in two States. A section of the scale is reproduced in Fig. 41. The numbers at the top of the columns represent the per cent of correct answers which were given by the 1696 pupils. These per cents are tentative standards.

Using the scale. This geography scale is a classified list of questions from which the teacher can select questions for a test. Since the questions in any column are equally difficult, it is best to take the questions for a test from one column. These may be given in the usual way by writing them on the board. It is better if each pupil is provided with a mimeographed copy with space left for writing in the answers. The teacher should not explain the meaning of any words used in the questions because the results will then not be comparable with the standards. Ten questions will make a test of convenient length and probably is the best number to use.

Scoring the papers. When ten questions are used and all have been chosen from the same column, each one may be considered to have a value of 10 credits. This will make the total number of credits 100. For scoring the papers the authors have prepared a score card. The portion of it which applies to the section of the scale reproduced in Fig. 41, is given below. These directions must be followed if the resulting scores are compared with the standards. The score of a pupil is the sum of the credits which he earns on the list of questions. The class score is the average of the scores of the members of the class.

What to accept and what to reject in scoring answers. Many of the exercises in the Hahn-Lackey Geography Scale admit of only one answer; but the scale contains exercises to which the answers may vary. In order that the scoring of the answers by teachers of different localities may be uniform and the scores be comparable with those of the scale, the authors have prepared a list of typical answers they accepted and typical answers they rejected in making the scale. In this list the answers to the exercises are given in the order in which they occur in the position of the scale reproduced in Fig. 41.

"F" means full credit; "P" means part credit; "N" means no credit.

42. Different kinds of meat was credited as only one kind of food.

1. F. "Brazil"; "South America"; "Mexico"; "Central America."

103. F. "Rays fall more vertical"; "Farther south"; "Nearer the equator."
N. "Ocean breezes"; "Gulf Stream."

104. F. "Capacity for water"; "Can go a long time without a drink."

18. F. "Let it down in the valleys or sea"; "Leave it on the bank"; "Form Islands"; "Drop it at the mouth"; "Drop it when the current is not swift." (Any answer that indicates that rivers take soil from one place and put it in another place.)

29. F. Any two lines of work; as, "Plants grain and harvests it"; "Raising cattle and farming"; "Raises crops and milks"; "Farming and selling things."
P. One half credit for only one line of work.

98. F. "No food for horses"; "Too cold for horses."

6. F. "Earth" "Land" "Rock" "Stone" "Mountains"; "Sand"; "Plains"; "Plateaus." (Any answer indicating a knowledge of the bed of the ocean.)

Q	R	S	T
42	50	58	66
58	66	73	79
66	73	79	84
79	84	88	92
79	84	88	92
60. Name four fur-bearing animals. 86. Name four large rivers of Europe. 99. What two countries of Asia are noted for tea?	62. Name the five Great Lakes of North America. 73. What are the two largest cities of the United States? 29. Name two ways in	52. What is the largest city of your state? 64. Where is Alaska and to whom does it belong? 84. Name four large cities of Europe.	35. Name two kinds of work that men do to keep us warm. 44. Name four things out of which clothing is made. 47. Give two ways in which

Colorado and Massachusetts.

42. Name two kinds of food that we get from animals.

1. From what country do we get much of our coffee ?

12. Which is the largest and which is the smallest, the moon, the sun, or the earth ?

71. What country of South America has a climate similar to ours?

103. Why is it warmer in New Orleans than in Chicago?

104. Why are camels such useful animals for traveling in desert places ?

9. What is the cause of day and night?

18. What do the rivers do with the soil that they carry?

41. Give two ways in which water gets away when it rains.

3. What is the name of the circle extending around the earth midway between the poles ?

98. Give one reason why the people in the far north use reindeer and dogs instead of horses.

37. Why does not the water flow out of swampy places?

46. Why do not the Eskimos build houses like ours?

6. What is under the ocean?

19. Where does the water in a well come from?

23. Name two ways in which winds are useful.

France and Germany.

101. Name two large bodies of water that border on Florida.

45. Name four things you use for food that do not grow where you live.

68. Give one reason why so many of the great cities of the United States are near the sea coast.

72. Which is the coldest and which the warmest part of South America?

48. Give two ways in which goods may be taken across a river.

7. Name four continents.

13. Name two necessary things that we get from the sun.

16. Name four races of men.

124. Give one way in which farming may be done in dry countries.

57. Give two ways in which rivers are useful.

38. Why should roads or streets be higher in the middle than on the sides?

39. How does the ocean help to furnish rain?

20. Where do plants get water?

21. What makes clouds move?

Fig. 41. A Section of the Hahn-Lackey Geography Scale.

19. F. "Out of the ground"; "From veins"; "Springs"; "From little streams in under the ground"; "From moisture in the ground." (Any answer clearly implying underground source.)
23. F. Any two. "Bring rain"; "Turn mills"; "Fresh air"; "Makes temperature equal"; "Makes pressure equal"; "Scatters seeds"; "Causes ocean currents."
68. F. "Good harbors"; "Commerce"; "Importing and exporting goods"; "Trade carried on there."
47. F. Any two. "By railroad"; "Wagon"; "Mail"; "Ships."
48. F. Any two. "Railroad bridge"; "Boat"; "Wagon bridge"; "Fording."
57. F. Any two. "Fish"; "Navigation"; "Sport"; "Water-power"; "Drains the land."
39. F. Any answer indicating a knowledge of the ocean as a source of vapor in the air; as, "Fills the air with vapor"; "Makes air moist."
21. F. "Winds"; "Breezes"; "Air"; "Atmosphere."

Interpreting the scores. Two things must be kept in mind if a satisfactory interpretation of the scores is obtained: (1) The standards given at the head of the columns do not represent the exact per cent of correct answers obtained. For example, 50, the fourth grade standard for column R, represents values from 46.8 to 53.2. Thus, small differences between scores and standards are not significant. (2) The standards are given in terms of a scale of 100 units or a per cent scale. It is also customary to use a per cent scale for reporting "school grades," or at least for defining them. Some per cent, as "60," "70," or "75," is chosen as the "passing mark." A score of 79 on a set of questions chosen from column S of this scale does not mean "79 per cent" as a school mark. This score of 79 is above standard in the fourth and fifth grades. It is just standard in the sixth grade and below standard in the seventh and eighth grades.

When a pupil has a score which is just standard, he is just an average pupil because the standard is an average score.

If five marks, such as A, B, C, D, and E (A being the highest mark), are used, such a pupil would be given a mark of C. If "grades" are reported in per cents and 75 is the passing mark, a score of 79 when the standard is 79 should be translated as about "85 per cent." Scores above and below standard should be interpreted with the standard representing the average.

II. American History

Harlan's Test of Information in American History. The items of information called for by this test are found in practically all American history textbooks, being based on the study by Bagley and Rugg of twenty-three textbooks to determine the content of American history.[1] This insures the questions being well selected. The credit to be given for answering each question has been determined and a score card is furnished with the test so that the marking of test papers may be uniform. The test consists of ten exercises. The first, fourth, and ninth are reproduced to illustrate their nature.

Exercise I

At the right of the page are the names of some men mentioned in American history. Fill in blanks with the names which properly belong there.

1. America was discovered by..............

Score. .... near the close of the fifteenth century.....

2. The name of the man who is supposed to have discovered the Pacific Ocean is......

3. The first President of the United States was ....

Jefferson
Cornwallis
William Penn
Lafayette
Patrick Henry

---

[1] Bagley, W. C., and Rugg, H. O., *The Content of American History as Taught in the Seventh and Eighth Grades.* (University of Illinois, School of Education, Bulletin no. 16, 1916.)

4. ....is the name of a distinguished Frenchman who aided the colonists in securing their independence.....

5. ....surrendered to the colonial troops at Yorktown............

Columbus
Benj. Franklin
Washington
John Cabot
Balboa

Exercise IV

Tell the very first thing you would do under each of the following conditions; also what you would do next:

1. If a neighbor were to present to you for your signature a petition to have some man removed from public office, —

Score......

What would you do first?...........................
Would you sign the petition?........................

2. If a man imprisoned in the county jail for some serious crime should be taken out by a mob with the intention of hanging him, —

What ought to be done first? ....
Then what?......................................

Exercise IX

The following topics represent matters of importance in the history of the United States. State definitely of what significance each has been.

1. Articles of Confederation..............................

Score...... ....

2. Mason and Dixon's line..............................
....

3. Monroe Doctrine..................................
....

4. The Tariff.......................................
....

Standards. The total amount of credit allowed for the ten exercises is 100. Over two thousand pupils in the seventh and eighth grades have been tested at the end of the respective years. The median scores are: Seventh grade, 56; eighth grade, 86.

As in the case of geography, these standard scores are the scores which should be made by the average or median (middle) pupils. Consequently, when translating these scores into school marks, a score which is standard should be given the "school grade" which represents the average pupil.

Diagnostic value of the test. The different exercises call for different types of information: dates, names, causes, meanings of historical terms, etc. By tabulating the credits earned on the different exercises in the way given for Charters's Diagnostic Tests in Language and Grammar (Fig. 39), a teacher may learn which phases of history need additional emphasis.

Other tests in geography and American history. Since tests in these two fields have been devised only recently and have not been used sufficiently to determine which ones are most helpful to the teacher we list here other tests. Information concerning them may be obtained by consulting the reference given or by writing to the address.

Geography

1. The Boston Tests. The two tests of this series — one on the United States and the other on Europe — consist of well-chosen questions. The relative difficulty of the questions was determined upon the basis of the per cent of correct answers. The tests were devised in an effort to determine (1) the character of achievement in geography, and (2) the possibility of scientific measurement of educational results in geography. This significant comment is made: "The results show how inadequate the customary examination or test in geography is to measure ability in geography."[1]

---

[1] *Geography; A Report on a Preliminary Attempt to Measure Some Educational Results.* (Boston, Department of Educational Investigation and Measurement, Bulletin no. 5, School Document no. 14, 1915; p. 38.)

2. Buckingham's Geography Test. This test was devised for use in the survey of the Gary and Prevocational Schools of New York City. It consists of two sets of twenty questions which were evaluated upon the basis of the per cent of correct responses.[1]

3. Starch's Geography Tests, Series A. The common elements of five geography texts have been arranged in five parallel tests. The exercises of the tests are in the form of mutilated sentences.[2]

4. Witham's Standard Geography Tests. These are a series of tests arranged to test quickly and easily the pupil's knowledge of certain geographical facts. The facts for the tests on the world are grouped under these heads: (1) geographical divisions, (2) form and motion of the earth, (3) the hemispheres, (4) land and water forms, (5) homes of the races, (6) industries, and (7) largest cities.[3]

5. Branom and Reavis Completion Test for the Measurement of Minimal Geographic Knowledge of Elementary School-Children. This test was very carefully constructed and appears to have a high degree of merit. No standards have been published.[4]

History

1. Buckingham's Tests. These tests were used in the survey of the Gary and Prevocational Schools of New York City. They consist of two sets of questions which have been evaluated on the basis of the per cent of correct answers.

---

[1] *Seventeenth Annual Report of the City Superintendent of Schools.* (New York City, 1914–15.)

[2] Address Daniel Starch, University of Wisconsin, Madison, Wisconsin.

[3] Witham, E. C., "A Minimum Standard for Measuring Geography"; in *American School Board Journal* (January, 1915), vol. I, pp. 13–14. Address E. C. Witham, Southington, Conn.

[4] *Seventeenth Yearbook of the National Society for the Study of Education,* part I, pp. 27–39. (Public School Publishing Company, Bloomington, Illinois, 1918.)

More recently Buckingham has studied the relation between the ability to remember historical facts and the ability to use them. In this study specially devised tests were used.[1]

2. The Bell and McCollum Test. This test consists of a series of questions which have been very carefully selected because of their importance. The topics included are: (1) dates-events, (2) men-events, (3) events-men, (4) historic terms, (5) political parties, (6) divisions of history, and (7) map-study. The test can be administered in a forty-minute period.[2]

3. Starch's American History Tests, Series A. This test is based upon the facts and principles common to five modern texts. The exercises are in the form of mutilated sentences. Four duplicate forms are available.[3]

QUESTIONS AND TOPICS FOR STUDY

1. What do you think of the method used in selecting the questions for the Hahn-Lackey Geography Scale? Is it a good method? Why?
2. Can we determine what pupils should know by examining textbooks? Why? In such a subject as geography, how can you determine what pupils should know?
3. What is a "tool subject"? A "content subject"?
4. How is a score on the Hahn-Lackey Geography Scale to be interpreted?
5. Select ten questions from the Hahn-Lackey Geography Scale. In what ways do they form a better test than a list of questions you might prepare?
6. How was Harlan's American History Test made? What are its strong points? Compare it with the usual history examination.

---

[1] *Seventeenth Annual Report of the City Superintendent of Schools.* (New York City, 1914–15.)

[2] Bell, J. C., and McCollum, D. F., "A Study of the Attainments of Pupils in United States History"; in *Journal of Educational Psychology* (May, 1917), vol. 8, pp. 257–74.

[3] Address Daniel Starch, University of Wisconsin, Madison, Wisconsin.

CHAPTER XI

EDUCATIONAL MEASUREMENTS AND THE TEACHER

In Chapter I evidence was presented which showed that measurements of the results of teaching, both by teachers' estimates and by examinations, were strikingly inaccurate. More accurate measurements of the results of teaching can be secured in two ways: (1) By using standardized tests. A number of these instruments have been described in the preceding chapters and the teacher has been told how to make the greatest use of the information obtained. (2) By introducing certain improvements into the making and using of examinations set by the teacher. Before presenting suggestions for making these improvements, it may be helpful to consider a question which may have occurred to some readers: *What is the value of measuring accurately the results of teaching?* The answer to this question will be presented under two heads: (1) A common-sense answer, or logical reasons; (2) experimental evidence.

(1) **Logical reasons for the value of accurate measurements by means of standardized tests.** It is a generally accepted principle that in any field of human endeavor the most efficient results are attained when one has a definite aim to work for and instruments for determining from time to time how much has been accomplished. A definite aim makes it possible for the worker to direct his efforts toward the particular things to be accomplished, instead of scattering them over an indefinite field. Instruments for measuring results make it possible for the worker to know what he has accomplished and where he needs to place greater effort and where less effort. It is also possible

for him to learn whether the materials and methods he is using are effective and which materials and methods are most effective.

This principle is the foundation of success in the business world. The merchant, manufacturer, or farmer who does not have a definite aim and who does not measure his accomplishment from time to time is most frequently a failure. Bankers measure the results of their business every day. Merchants take inventories once a year and in many cases more frequently, sometimes a partial inventory every day. They consider this procedure necessary for success.

The teacher is a manufacturer. His raw material is the children. Textbooks, school buildings, equipment, libraries, and methods and devices of teaching are the "machines" or instruments which he uses to change this raw material into the finished product or educated boys and girls who are prepared to do their part in the life of the community, State, and Nation. Without definite aims the teacher cannot plan his work effectively. He does not know, except in an indefinite or general way, what he is to do. If he has definite aims, but no instruments for measuring his results accurately, he cannot learn when he has attained his aims. Thus he is compelled to work in the dark. If he makes inaccurate measurements, but considers them accurate, he is in a still more serious situation. His efforts are almost certain to be expended unwisely.

Evidences of the lack of definite aims. It may appear to some teachers that the aims for the several grades as stated in courses of study are sufficiently definite. In the case of reading it was pointed out on pages 82–85, that oral reading was sometimes over-emphasized, while silent reading was neglected, and that in certain cases the rate of silent reading was not sufficiently emphasized. In the operations of arithmetic the rate of work is sometimes neglected (see pages 23 and 25). Cases of irregular progress are shown in Figs. 119 and 122. When classes make irregular progress in ability to do the different types of examples, as is shown in these figures, the lack of definite aims is a contributing cause.

Another type of evidence is given in Fig. 42.[1] This figure is drawn by the method given on page 110, and shows five sets of median scores for rate and accuracy in Addition (Series B) in Grades four to eight: (1) Courtis Medians, (2) Medians for Cuyahoga County (Ohio), (3) Brook Park, (4) Rocky River, and (5) Shaker Heights. (The last three are three district schools in Cuyahoga County.) The broken line (----) connecting the small circles representing the Courtis Medians shows gradual and regular progress from grade to grade. The same condition is exhibited for Cuyahoga County as a whole, but a distinctly different situation is shown for the three schools. In the fourth grade the Brook Park School is very low in accuracy. The fifth-grade pupils show a decided gain in accuracy, but almost no progress in rate of work, while the accuracy median for the sixth grade is below that for the fifth grade. The rate of work for these three grades is approximately the same. From the sixth grade to the seventh marked progress in both rate and accuracy are shown and from the seventh to the eighth grade the progress is almost wholly in accuracy. A study of the lines representing the median scores of Rocky River and Shaker Heights reveals similar irregularities in the progress from grade to grade. This condition is probably due in a large degree to the fact that the teachers did not have definite aims for the several grades and had not

---

[1] This figure is taken from the *Third Annual Bulletin and School Directory of Cuyahoga County (Ohio) School District*, p. 71, by County Superintendent A. G. Yawberg.

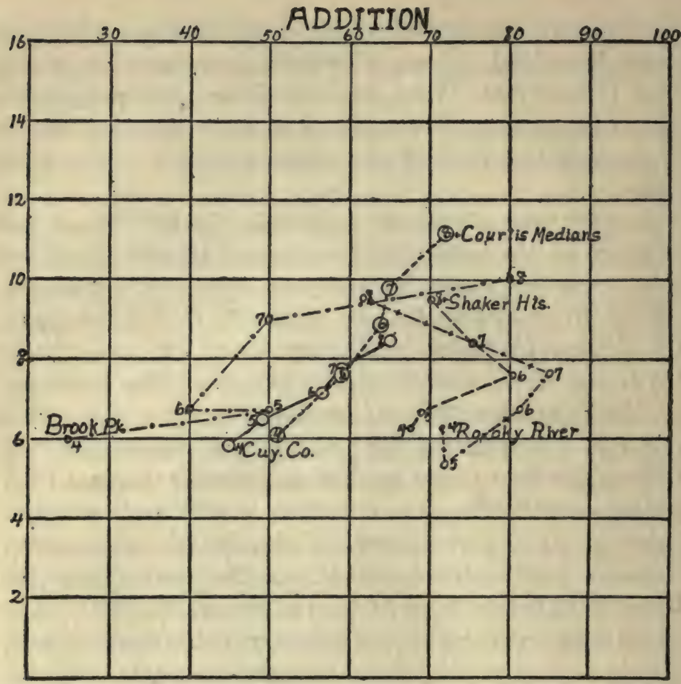


FIG. 42. MEDIANS IN SPEED AND ACCURACY IN ADDITION TEST, SERIES B, FOR PUPILS OF GRADES 4 TO 8 INCLUSIVE, SHOWING COURTIS'S MEDIANS, MEDIANS FOR CUYAHOGA COUNTY, AND FOR THREE DISTRICTS IN THE COUNTY. RATE ON VERTICAL SCALE, ACCURACY ON HORIZONTAL SCALE.

measured the results of their instruction to learn to what extent these aims were being realized.

Courses of study do not give definite working specifications for teachers. Our present courses of study represent the efforts of those occupying supervisory positions in our school systems to provide teachers with specifications for their work. How well they have succeeded is illustrated by the following quotations from typical courses of study:

Fourth Grade

Reading and Literature. Stories read and told to the class; Roman stories, American history stories relating to geography, selections from Greek and Teutonic mythology, and poems.

A few choice selections of appropriate prose and poetry are to be studied, committed to memory, and recited or dramatized. See that the pupil stands on both feet and reads smoothly and confidently. Watch the voice of pupils; use breathing exercises; and avoid harsh, strained reading. Have pupils read many selections silently, then reproduce the thought aloud in order to develop the power of gaining and expressing the thought of the text. Aim to enlarge the pupil's vocabulary, to help him master the thought content, and gain the power to read in a pleasing, well-modulated tone. Use much supplemental reading. Explain the purpose of the children's department in the Public Library, and encourage pupils to read books therefrom.

Following these general directions, the selections to be read are specified:

Third Grade : B Class

Handwriting. Daily drill, lessons five, six, or seven.[1] During the entire year these drills should be used for a few moments at the beginning of each writing lesson.

Beginning with lesson five, take the lessons in consecutive order to lesson thirty-five.

After developing a letter with the class take, from the writing book, a word beginning with the same letter and use it for practice.

If the letter is a capital, follow the word practice by using a sentence beginning with this letter. All words and sentences should be taken from the writing book.

Grade 4B. Arithmetic

Leading topics. The four fundamental processes with emphasis upon multiplication.

Review. Regularly, constantly, and from the first. The addition combinations, subtraction, reading and writing numbers, simple fractions.

---

[1] Reference to a *Manual for Teachers,* used in this school system.

Multiplication. The tables completed and made automatic. Problems with two-place multipliers. Rapid oral practice.

Division. Short division with long division brace. Rapid oral practice.

Fractions. Simple fractions and mixed numbers as needed in actual practice on concrete form problems. Largely oral and objective.

Concrete problems. One-step problems.

Applied problems. Farm products, farm marketing, farm profits.

Measures. Quart, gallon, peck, bushel, pound, ton, cord, etc., as required by applied problems.

In the fourth-grade course of study for reading, nothing is said about the rate at which a pupil is expected to read orally or the rate and degree of comprehension of silent reading. The teacher is told to "enlarge the pupils' vocabulary," but he is not told *how much* to enlarge it. In handwriting nothing is said about the rate or quality of writing. In arithmetic the multiplication tables are to be "completed and made automatic." This is indefinite because there are many degrees of automatization. What one teacher would call automatic, another would not. "Rapid oral practice" is likewise indefinite.

Definite specifications for these subjects could be given by stating the rate and quality of work expected of the pupils. For example, in handwriting the teacher would have a definite aim if he was told that fourth-grade pupils should be able to write memorized material at the rate of 56 letters per minute and with a quality of 50 as measured on the Ayres Scale. In arithmetic the aim would be definite if he were told that fourth-grade pupils should be able to write the answers to the multiplication combinations at the rate of 23 per minute. When the teacher has his aims expressed in this definite way, he still needs to measure results to learn to what extent he is realizing his aim.

(2) Experimental evidence of the value of standardized

tests to the teacher. Several instances of the value of standardized tests have been given in the preceding chapters. For reading, see pages 73 to 89; for arithmetic, see page 151; additional evidence is given in Figs. 43, 44, and 45. In Fig. 43 a sixth-grade teacher in a small town gave the Courtis Arithmetic Tests, Series B, in September, 1917. The condition revealed is shown by the lower line in the figure. The standards or the definite aims to be attained are represented by the heavy solid line. The results of the teacher's efforts are shown by the upper line which represents the median scores for April, 1918. This teacher knew what her pupils could do at the beginning

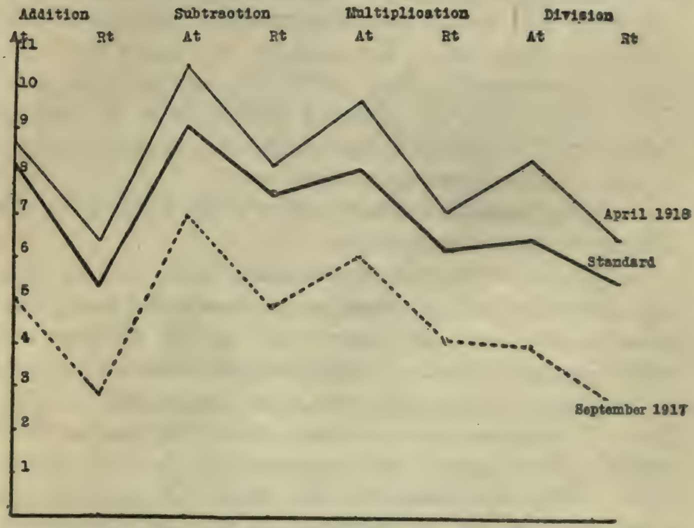


Fig 43. Showing the Median Scores of a Sixth-Grade Class in September, 1917, and in April, 1918, as measured by the Courtis Standard Research Tests in Arithmetic, Series B

Dotted line September medians. Solid line (upper) April medians. Heavy solid line, standard median scores.

of the year and the standards which were to be attained. She had the satisfaction of knowing that these had been attained.

In Cleveland, Ohio, spelling was tested in May, 1915. It was again tested in May, 1917. In Fig. 44 the "efficiency" of spelling in certain of the buildings is showing for these two tests. The broken line represents the "efficiency"[1] for the test in May, 1915. The solid line represents the "efficiency" for the test in May, 1917. The significant thing in this illustration is the material improvement in spelling "efficiency" of those buildings which had low average scores in 1915. This suggests that when a teacher knows that he is not securing standard results, he is likely to increase the effectiveness of his instruction.

In the city of Boston standardized tests in arithmetic were given in 1912 in twenty-nine schools and have been given at least once each year since. In Fig. 45 these are called Group A Schools. In the Group B Schools the standardized tests in arithmetic were given first in 1913 or 1914. Group C includes those schools in which the results of instruction in arithmetic were measured for the first time in 1915. Group A represents 18,391 pupils, Group B, 15,241 pupils, and Group C, 11,836 pupils. In May, 1915, standardized tests in arithmetic (the Courtis, Series B) had been used in Group A for three consecutive years, in Group B for one or two years, and in Group C they were being given for the first time.

We have here, therefore, an opportunity to study the effect of using such tests. If the tests are helpful to the teacher, those schools in which they have been used for three

---

[1] "Efficiency" means here the quotient of the average score divided by the standard given by Ayres on his Spelling Scale. Since these standards are lower than the average scores in certain cases those quotients are above 100.

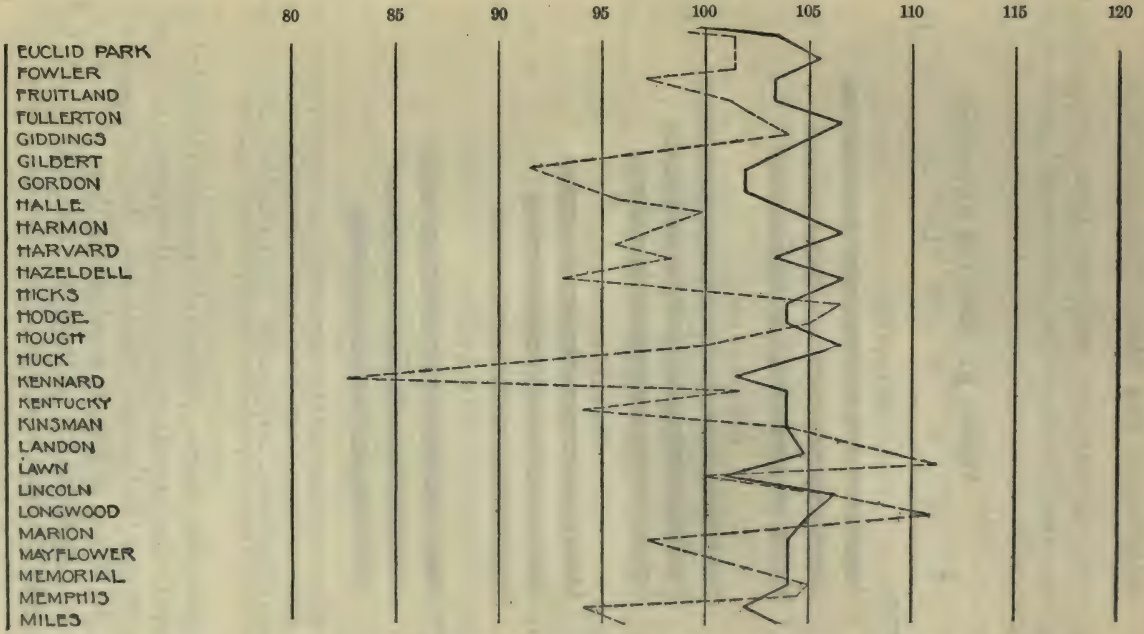


FIG. 44. SHOWING IMPROVEMENT IN SPELLING "EFFICIENCY" IN CERTAIN SCHOOLS OF CLEVELAND, OHIO

Taken from *The Results of a Spelling Test*, Bulletin no. 2, Board of Education, Division of Reference and Research, March, 1918.

years should show the best results and those in which they have been used for one or two years should stand above those in which they were given for the first time. Fig. 45 represents the median number of examples attempted by the eighth-grade pupils in each of the three groups of schools. In every case the Group A Schools have the highest medians and the Group B Schools stand above those of Group C. The median scores for accuracy are not represented, but they show the same condition. Thus, this figure shows that

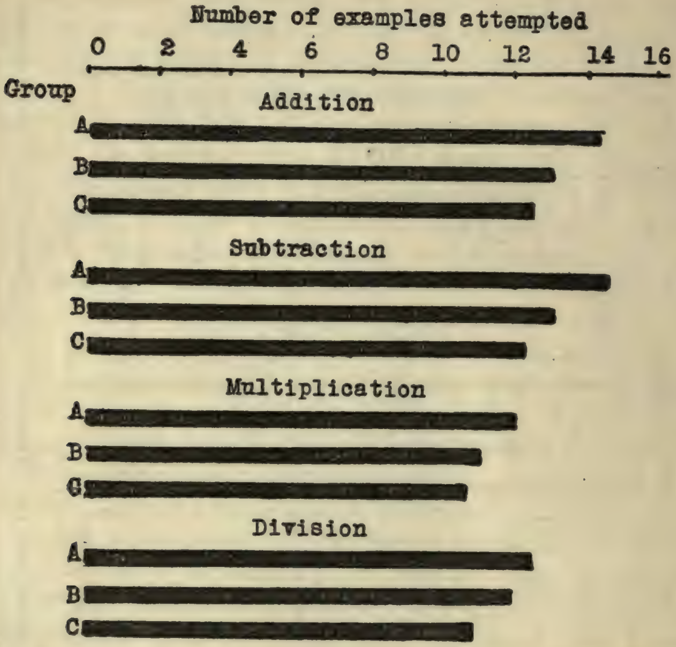


FIG. 45. SHOWING EFFECT OF CONTINUOUS USE OF COURTIS'S STANDARD RESEARCH TESTS, SERIES B, IN BOSTON, EIGHTH GRADE, 1915

Group A schools continuous use for three years. Group B schools use for one to two years. Group C schools not given until this record was secured. (*After Ballou.*)

in the case of the eighth grade superior results in the operations of arithmetic are attained by those schools in which standardized tests are used. The results for the other grades in Boston are not so striking, but they furnish additional evidence that it is helpful to measure the results of instruction accurately.

Making examinations yield more accurate measurements. In Chapter I the following criticisms were made of the measurement of the results of teaching by examinations: (1) The questions cover a wide range of topics making the "grade" have no definite meaning. (2) The questions are generally not equally difficult and the same amount of credit should not be given for answering different questions. The judgment of a teacher in regard to the amount of credit which should be given for answering a question correctly is not reliable. (3) Teachers do not mark examination papers accurately. (4) The pupil's rate of work is usually neglected even in those subjects where it is important. (5) Standards are not available for interpreting the measures.

It is not possible for the teacher to eliminate entirely the defects enumerated by these criticisms, but it is possible for him to reduce them. Greater care in choosing and framing the questions will materially reduce the first defect mentioned above. Catch questions should always be avoided. Also the question should be stated so that the pupil will understand what is called for. The questions should be *important.* Unimportant facts should not be called for unless there is some particular reason for doing so.

The amount of credit to be given for answering a question correctly cannot be accurately determined, but a helpful rule to follow for questions which are approximately equal in importance is that the most credit should be given for the most difficult question and the least credit for the easiest.

Both of the above suggestions will tend to increase the accuracy of the marking of examination papers. In addition a systematic plan will materially reduce this source of error. Kelly[1] describes the following experiment: Six fifth-grade teachers gave a uniform examination in arithmetic to their pupils. Each teacher marked the papers for her own pupils, but did not record the marks on the papers. The superintendent asked a teacher, who was unusually systematic in marking examination papers, to prepare a definite plan for marking these papers. After she had done so, she marked all of the papers in accordance with this plan. Then the teachers who had first marked the papers marked them a second time following her plan. This provided two marks for each paper by the classroom teacher, the first without following a systematic plan, and the second using a definite plan. Each of these marks was compared with the mark of the teacher who marked all of the papers. In Table XXIV the six teachers are designated by the letters A, B, C, D, E, and F. The table is read as follows: When no systematic plan was followed, teacher A marked one paper 16 to 20 points lower than the "judge," one paper 7 points lower, two papers 4 points lower, two papers 2 points lower, agreed with the "judge" on one paper, etc. The differences between the marks given when the classroom teachers had no systematic plan and when they followed such a plan are very striking. In the first instance the marks assigned by the teachers agreed with those assigned by the "judge" in only 5.5 per cent of the cases, while in the second instance they agreed in 63.5 per cent of the cases. Thus the experiment shows that with a systematic plan for marking papers, the marks will be more accurate.

The rate at which the pupil works can easily be measured in such subjects as reading, handwriting, and the operations of arithmetic. It is only necessary to place a time limit upon the examination such that no pupil will answer all of the questions. The number of questions answered will be a crude measure of his rate of work.

---

[1] Kelly, F. J., *Teachers' Marks*, p. 84.

TABLE XXIV. DISTRIBUTIONS OF DIFFERENCES BETWEEN TWO SETS OF TEACHERS' MARKS ON FIFTH-GRADE ARITHMETIC PAPERS — FIRST, WITHOUT ANY EFFORT TO UNIFY THE METHODS USED, AND SECOND, BY A COMMON STANDARD (AFTER KELLY)

Range of Differences	*Without standard*							*With standard*						
	A	*B*	*C*	*D*	*E*	*F*	*Total*	*A*	*B*	*C*	*D*	*E*	*F*	*Total*
21 or more	..	..	..	..	2	..	2	..	..	..	..	..	..	..
16 to 20	1	..	..	..	1	1	3	..	..	..	..	..	..	..
15	..	..	..	..	..	2	2	..	..	..	..	..	..	..
14	..	..	..	..	..	1	1	..	..	..	..	..	..	..
13	..	..	..	..	1	2	3	..	..	..	..	..	..	..
12	..	1	..	..	1	..	2	..	..	..	..	..	..	..
11	..	..	1	..	1	2	4	..	..	1	..	..	..	1
10	..	..	..	..	..	1	1	..	1	..	..	..	..	1
9	..	1	..	..	2	1	4	..	..	..	..	..	..	..
8	..	..	..	1	3	1	5	..	..	..	..	..	..	..
7	1	1	..	1	1	1	5	..	..	..	1	..	..	1
6	..	2	..	..	1	1	4	..	..	..	..	..	..	..
5	..	1	2	1	1	2	7	..	..	..	..	..	..	..
4	2	2	2	1	1	2	10	1	..	..	..	..	1	2
3	..	4	2	1	2	2	11	..	..	1	1	1	..	3
2	2	2	1	1	1	1	8	4	1	1	3	7	1	17
1	..	5	4	3	2	4	18	2	3	4	5	1	1	16
0	1	4	4	1	1	1	12	22	30	16	16	29	26	139
1	2	5	2	2	2	1	14	5	..	2	2	1	3	13
2	6	1	3	2	3	1	16	1	1	3	..	..	..	5
3	9	..	2	..	2	..	13	..	2	2	1	..	1	6
4	5	1	4	1	5	1	17	..	2	3	3	..	..	8
5	2	3	2	2	1	..	10	..	1	1	2	..	..	4
6	1	1	..	3	2	..	7	..	..	1	1	..	..	2
7	..	1	1	6	1	..	9	..	..	..	..	..	..	..
8	..	2	1	2	..	1	6	..	..	..	..	..	..	..
9	1	..	1	2	..	..	4	..	..	..	..	..	..	..
10	1	..	..	1	..	1	3	..	..	..	..	..	..	..
11	..	1	1	..	..	..	2	..	..	..	..	..	..	..
12	1	..	..	1	..	1	3	..	.	..	..	..	..	..
13	..	1	..	..	1	1	3	..	..	..	..	..	..	..
14	..	..	..	..	..	1	1	..	..	..	..	..	..	..
15	..	..	1	1	..	..	2	..	..	..	1	..	..	1
16 to 20	..	2	..	..	..	..	2	..	..	..	..	..	..	..
21 or more	..	..	1	3	1	..	5	..	..	..	..	..	..	..
Totals	35	41	35	36	39	33	219	35	41	35	36	39	33	219
Medians	+3	0	+1	+6	—1	—4	+1							

The lack of standards can be partially remedied by having other teachers give the same examination to other pupils of the same grade. Where this is not possible, the teacher should verify what he considers a satisfactory standard by using standardized tests occasionally. If the standardized tests show a class to be near the standard, the teacher may conclude that his standard is satisfactory. If the work of the class is shown to be unsatisfactory then the teacher should conclude that his standards are too low unless her examinations have also shown the class to be doing a low grade of work.

QUESTIONS AND TOPICS FOR STUDY

1. How may examinations be made more accurate measuring instruments?
2. Do you think that examinations have functions other than that of measuring the abilities of pupils? If so, what are they?
3. Repeat the experiment described on page 218 and compare your results with those given in Table XXIV.
4. Why is a general aim not sufficient?
5. What are the objections to our present courses of study with respect to the statement of aim?
6. How can standardized tests be used in setting the aim for the teacher?
7. Criticize your course of study. How could you make use of standardized tests in improving it?
8. Why should teachers use standardized tests? (Give all of the reasons you can think of.) Do they require more time on the part of the teacher than similar tests he might prepare?

CHAPTER XII

SUMMARY

The use of standardized tests by teachers may be summarized under the following steps.

1. Selection of a test to use. For the most part teachers should accept the advice of experts in selecting a test. However, it is well for teachers to ask these questions about a test: (1) Has it been widely used or is it likely to be widely used in the near future? (2) How much time is required to give the test, to mark the papers, and to record the scores? These points are very important. If the test has not been widely used or there is no prospect of its wide use in the near future, reliable standards will not be available. Also the general use of a test indicates that it has been found helpful to teachers. If a test requires a large expenditure of time, a teacher is not likely to receive adequate returns for the time spent. In general a teacher should choose a test which is simple to use and which requires only a moderate amount of time.

2. Giving the test. In giving a test to a class the teacher should follow the directions. If this is not done, comparisons of the resulting scores with the standards will not be valid. Also the teacher should bear in mind that his purpose should not be to secure as high scores as possible, but to secure a *true* measure of the abilities of his pupils. In order to do this the pupils must not be excited or urged to work in an unnatural way.

The manner in which the test is presented to the pupils affects the scores. The purpose of measurement is defeated if the test is presented to the pupils in such a way that their response is unnatural. For example, in handwriting, if the pupils write at an unnatural rate the quality of their handwriting will be affected.

3. **Marking the test papers.** Explicit directions and score cards are generally provided for doing this. Here also the teacher must follow the directions. If he does not, a valid comparison cannot be made with the standards. In the case of a few tests, the papers are marked by the pupils as the teacher reads the correct answers. This can be done in the case of tests in the operations of arithmetic, in certain reading tests, and in spelling. Courtis advises that this plan be followed in order to save time which may be used in the interpretation of scores. In the case of handwriting and composition, the teacher should equip himself for accurate rating of papers by systematic training.

4. **Recording the scores and calculating class scores.** Blanks for recording the scores are usually furnished with the tests. When they are at hand, this step is very simple after the teacher has had a little practice.

5. **The interpretation of scores.** In interpreting both individual and class scores, standards are necessary. Many people understand facts more easily when they are represented graphically. Hence, it is well to employ some means of graphical representation.

6. **Correction of the defects revealed by the test.** The correction of the defects revealed by the test is the culmination of the preceding steps. It is in this step that the value of standardized tests is realized. Without this step standardized tests become mere "playthings" and their use cannot be justified. The situation created is similar to that which would exist if a physician examined a patient carefully and determined the nature of his ailment, but did not prescribe any remedial treatment. In our zeal to convert teachers to the acceptance of the principle that the measurement of

certain results of instruction is possible, there has been a tendency to overlook this step. In fact some have even said that they were content to apply the tests and reveal to the teachers the shortcomings of their work. These persons would leave to the teachers the difficult problem of remedying the defects. As a result not a few teachers have failed to see in the tests anything more than a new "plaything," which they might use to secure material for a paper to read at a teachers' association or to arouse the interest of their pupils. Such teachers have expressed their approval of the tests when their pupils' scores were high, and have considered the tests unsatisfactory when the scores were low.

In order to prescribe the best corrective instruction the teacher needs to have as much information as possible. Hence, the need for diagnostic tests and for examining the test papers to learn of the pupils' errors. Diagnosis requires time, but it is justified by making possible the planning of more effective corrective instruction.

APPENDIX

A sample package of tests. It will be helpful to a teacher in reading this book to have at hand sample copies of the tests described in it. In a few cases it is almost necessary to have a copy of the test in order to understand the discussion of its use. Believing that teachers would appreciate the opportunity of being able to secure all of the tests from one address, the author has assembled packages containing one copy of the tests marked with a star below. The other tests have been either reproduced in the pages of this book or described so fully that a copy is not needed in order to understand the test. A sample package will be sent postpaid upon receipt of a post-office money order or a check for 40 cents. Do not send stamps. Coins may be sent at sender's risk. Address Walter S. Monroe, Bureau of Coöperative Research, Indiana University, Bloomington, Indiana.

Ordering tests for class use. All of the tests listed in the following table can be obtained from the various publishers. A large number of them can be obtained from the distributing centers listed below. In ordering from one of these centers there is the advantage of being able to secure all, or at least several, of the tests desired from one address. In addition these Bureaus are prepared to render other important service. Hence, it is recommended that teachers send their orders to the Bureau of the State in which they reside. If their State has no Bureau, they may send their orders to the nearest one. The prices of the tests when ordered from a Bureau are generally the same as when ordered from the publisher.

DISTRIBUTING CENTERS

Bureau of Educational Research, University of Illinois, Urbana, Ill.

Bureau of Coöperative Research, University of Indiana, Bloomington, Indiana.

Educational Extension Service, University of Iowa, Iowa City, Iowa.

Bureau of Educational Measurements and Standards, Kansas State Normal School, Emporia, Kansas.

Bureau of Coöperative Research, University of Minnesota, Minneapolis, Minnesota.

In the table below detailed information about the tests described in the preceding chapters is given. One not familiar with ordering tests should study this table carefully in order to be able to ask for just what is needed. In some cases the necessary directions and record sheets are furnished by the publisher, but this is not true in all cases. When it is not done, the one ordering must ask for the number of these accessories desired.

Very important. In ordering tests of which one copy is needed for each pupil always give the number of pupils in each grade. This is important because some of the series have different tests for the different grades.

Prices. The prices given below are subject to change. In some cases the amount of postage is given. In practically all other cases the purchaser is charged with the postage but the author does not have information in regard to the amount.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
*1. Monroe's Standardized Silent Reading Tests.	One copy of the test for each pupil.	All directions are printed on either the test or on the class record sheet. One record sheet is furnished with each 25 copies of the test. Additional copies may be ordered if desired.	3 to 8	22	Bureau of Educational Measurements and Standards, Emporia, Kansas.	Including complete directions and record sheets, 60 cents per one hundred copies; postage extra, 9 cents per 100.
*2. Courtis's Silent Reading Test No. 2.	One copy of the test for each pupil.	Folder B, Series R, contains detailed directions for giving the test and for scoring by the pupils. One copy is needed for each person giving the test. Folder D, Series R, contains detailed directions for completing the scoring, for recording the scores and for calculating class scores. One copy is needed for each person giving the test. A class record sheet for recording the scores of a class is needed for each class. A school record and graph sheet for Silent Reading No. 2 is needed for each school.	2 to 6	29	S. A. Courtis, 82 Eliot St., Detroit, Mich.	Test only $1.80 per hundred; Folder B, 5 cents; Folder D, 5 cents; Class Record Sheet, 1½ cents each; Record Sheet No. 3 and Graph Sheet, 1½ cents each.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS (*continued*)

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
*3. Thorndike's Visual Vocabulary Test.	One copy of the test for each pupil.	As published by the Bureau of Educational Measurements and Standards, Emporia, Kansas, all directions are printed upon the test or class record sheet. A copy of the class record sheet is furnished with each 25 copies of the test. Additional copies may be ordered if desired. As published by the Bureau of Publications, Teachers College, New York City, detailed directions are not furnished. Also a preliminary test and a class record sheet must be ordered separately.	3 to 8	36	Bureau of Educational Measurements and Standards, Emporia, Kansas; or Bureau of Publications, Teachers College, New York City.	Including complete directions, preliminary test and record sheets, 50 cents per hundred; postage extra, 9 cents per hundred. Test only, 40 cents per hundred
*4. Gray's Oral Reading Test.	One copy of the test for each pupil and an extra copy for the teacher.	Directions for giving the test, for scoring the test papers, for recording the scores and for calculating class scores are printed on the class record sheet. One class record sheet should be ordered for each class.	2 to 8	39	W. S. Gray, University of Chicago, Chicago, Ill.	Including complete directions and record sheet, $1. per hundred.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS (*continued*)

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
*5. Courtis's Standard Research Tests in Arithmetic, Series B.	One copy of the test for each pupil.	Folder B, Series B, contains detailed directions for giving the tests. One copy is needed for each person giving the tests. Directions for recording the scores are found on the class record sheet. One copy is needed for each class. Folder D, Series B, contains directions for calculating class scores. One copy is needed for each person giving the tests. One record sheet No. 3, and graph sheet are needed for each school. One answer card is needed for each person giving the tests.	4 to 8	97	S. A. Courtis, 82 Eliot St,, Detroit, Mich.	Test only 80 cents per hundred ; Folder B, 5 cents; Folder D, 5 cents; Class Record Sheet, 1½ cents each. Record Sheet No. 3 and Graph Sheet, 1½ cents each. Answer card, 1½ cents each.
*6. Monroe's Diagnostic Tests in Arithmetic.	One copy of the test for each pupil.	Directions for giving the tests are printed on the tests and directions for handling the scores are printed on the class record sheet. One record sheet is furnished with each 25 tests. Additional class record sheets may be ordered if desired. One an-	Tests 1 to 11 in grades 4 to 8; Tests 12 to 16 in grades 5 to 8;	109	Bureau of Educational Measurements and Standards, Emporia, Kansas.	Including complete directions and record sheets : Tests 1 to 11, $1.20 per 100 ; postage extra, 20 cents per 100; Tests 12 to 16, 60 cents per 100 ; Tests 17 to

DIRECTIONS FOR ORDERING STANDARDIZED TESTS (*continued*)

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
		swer card is furnished with each 50 tests.	Tests 17 to 21 in grades 6 to 8.			21, 60 cents per 100; postage, extra, 10 cents per 100.
*7. Monroe's Standardized Reasoning Tests.	One copy of the test for each pupil.	Directions for giving the test are printed on the test. Directions for marking the papers and for recording the scores are printed on the class record sheet. One copy of the record sheet is furnished with each 25 tests. Additional Class Record Sheets may be ordered if desired.	4 to 8	154	Bureau of Coöperative Research, Indiana University, Bloomington, Ind.	Including complete directions and record sheets, 60 cents per 100; postage extra, 9 cents per 100.
*8. Ayres's Spelling Scale.	One copy of the scale for each teacher.	Brief directions for using the scale are printed on it. (See page 184 of this book for additional directions.)	2 to 8	176	Russell Sage Foundation, New York City.	Five cents per copy.
9. Timed-Sentence Spelling Tests.	One copy of the tests for each teacher.	Directions for using the test are printed on the test. Directions for handling the scores are printed on the class record sheet. One record sheet is furnished with each test. Addi-	3 to 8	184	Bureau of Coöperative Research, Indiana University, Bloomington, Ind.	Single sets, including record sheets, 10 cents each; three or more copies, 7 cents each.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS *(continued)*

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
		tional copies may be ordered if desired.				
*10, Ayres's Handwriting Scale, Gettysburg Edition.	One copy of the scale for each teacher.	Directions for using the scale are printed on the scale. No class record sheet is furnished by the publisher. Some of the distributing Bureaus named above have devised a class record sheet which they furnish with the scale.	2 to 8	208	Russell Sage Foundation, New York City.	Five cents per copy.
11. Gray's Score Card for Handwriting.	One copy of the score card for each pupil.		2 to 8	216	C. T. Gray, University of Texas, Austin, Texas.	
12. Freeman's Handwriting Scale.	One copy of the scale for each teacher.	No directions for using the scale are furnished by the publisher. (See page 215.)	2 to 8	212	Houghton Mifflin Company, Boston, Mass.	Twenty-five cents per copy.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS *(continued)*

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
13. Willing's Composition Scale for Measuring Written Composition.	One copy of the scale for each teacher.	Complete directions for using the scale are printed on it. One copy of the class record sheet is furnished with the test. Additional copies may be ordered if desired.	4 to 8	235	Bureau of Coöperative Research, Indiana University, Bloomington, Ind.	Single copies, 8 cents per copy. Five or more, 5 cents per copy; postage extra, 1 cent per copy.
14. Nassau County Supplement to the Hillegas Scale.	One copy of the scale for each teacher.	No directions for using the scale are furnished with the scale by the publishers. There is also no class record sheet.	4 to 8	243	Bureau of Publications, Teachers College, New York City.	Eight cents per copy by mail; in quantity, 5 cents per copy, postage extra.
*15. Charters's Diagnostic Test in Language and Grammar for Pronouns.	One copy of the test for each pupil.	Directions for giving the test are printed on the test. Directions for marking the papers and for handling the scores are printed on the class record sheet. One class record sheet is furnished with every 25 copies of the test. Additional copies may be ordered if desired.	7 and 8	245	Bureau of Educational Research, University of Illinois, Urbana, Ill.	Including complete directions and record sheet, 70 cents per 100.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS *(continued)*

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
*15a. Charters's Diagnostic Language Test	One copy of the test for each pupil.	Directions for giving the test are printed on the test. Directions for marking the papers and for handling the scores are printed on the class record sheet. One class record sheet is furnished with every 25 copies of the test. Additional copies may be ordered if desired.	3 to 8	245	Bureau of Educational Research, University of Illinois, Urbana, Ill.	Including complete directions and record sheet, 45 cents per 100.
16. Starch's Punctuation Scale.	One copy of the scale for each pupil.	Directions are given on a separate sheet. No class record sheet is furnished by the publisher.	5 to 8	248	Daniel Starch, University of Wisconsin, Madison, Wis.	Test, 45 cents per 100; directions, one cent per copy.
*17. Courtis's Standard Tests in Geography for States and important places.	One copy of the test for each pupil.		4 to 8	256	S. A. Courtis, 82 Eliot St., Detroit, Mich.	Instructions and tests for 40 pupils, 75 cents.

DIRECTIONS FOR ORDERING STANDARDIZED TESTS (*continued*)

Test	*How many tests to order*	*How many directions, record sheets, and other accessories to order*	*Used in what grades*	*Described on page of this book*	*Publisher*	*Price*
18. Hahn-Lackey Geography Scale.	One copy of the scale for each teacher.	Brief directions for using the scale are printed on it. Explicit directions for scoring the test papers are printed separately and furnished with the scale. There is no class record sheet.	3 to 8	258	H. H. Hahn, State Normal School, Wayne, Neb.	Seven cents per copy.
*19. Harlan's Test of Information in American History.	One copy of the test for each pupil.	Directions for giving the test, for scoring the papers and for handling the scores are printed separately. One copy is needed for each person giving the test. Class Record Sheet, one copy is needed for each class.	7 and 8	262	Bureau of Coöperative Research, University of Minnesota, Minneapolis, Minn.	Test only, 3 cents per copy; directions, one cent per copy; Class Record Sheet, one cent per copy.

INDEX

Adams, John, 170.
Analytical diagnosis in arithmetic, 138.
Arithmetic: Courtis Standard Research Tests, Series B, 97 ff., 119 ff.; Monroe's Diagnostic Tests, 109, 131 ff.; Cleveland Survey Tests, 151; Monroe's Standardized Reasoning Tests, 154 ff.; Stone's Reasoning Test, 173; Starch's Arithmetical Scale A, 174.
Arithmetic, types of examples, 111, 115.
Arithmetic, vocabulary in, 163, 165.
Ashbaugh, E. J., 129, 221.
Ayres, L. P., 176, 178, 208, 221.
Ayres's Handwriting Scale, "Gettysburg Edition," 207, 209 ff.; Ayres's Handwriting Scale, "Three-Slant Edition," 208.
Ayres's Measuring Scale for Ability in Spelling, 175 ff., 187–89.

Bagley, W. C., 262.
Ballou, F. W., 130.
Bell and McCollum History Test, 266.
Boston Copying Test, 249.
Boston Geography Tests, 264.
Branom and Reavis Completion Test for Geography, 265.
Breed, F. S., 209, 244.
Breed and Frostic Composition Scale, 244.
Brownell, Baker, 253.
Buckingham, B. R., 10, 256.
Buckingham's Geography Test, 265.
Buckingham's History Test, 265.
Charters, W. W., 230, 254.
Charters's Diagnostic Test in Language and Grammar, 245 ff.
Chase, Sara E., 165.
Cleveland Survey Tests, 151.
Comin, Robert, 11.
Composition: Willing's Scale for Measuring Written Composition, 235 ff.; Nassau County Supplement, 243; Hillegas Composition Scale, 244; Thorndike Extension of the Hillegas Scale, 244; Harvard-Newton Composition Scale, 244; Breed and Frostic Composition Scale, 244.
Corrective instruction: in reading, 58 ff., 65 ff., 72 ff., 85, 86 ff.; in arithmetic, 121, 123, 124, 128, 131, 135, 139 ff., 145 ff., 158–60, 168, 173; in spelling, 192 ff.; in handwriting, 224 ff., 228, 231 ff.
Counts, George S., 144.
Courtis, S. A., 107, 111, 138 ff., 182, 193.
Courtis's Silent Reading Test No. 2, 29 ff., 46 ff., 82 ff.
Courtis's Standard Practice Tests, 136.
Courtis's Standard Research Tests, Series B, 97 ff.; limitations of, 114.
Courtis's Standard Tests in Geography, 256 ff.

Diagnosis: in arithmetic, 119, 122, 123, 124 ff., 128 ff., 138 ff., 157 ff., 160 ff., 168 ff.; in reading, 53 ff., 65, 69 ff., 82 ff.; in spelling, 190 ff.; in handwriting, 212 ff., 231 ff.; in language and grammar, 245; in history, 264.

Educational measurements, value of, 267 ff.
Elliott, E. C., 8.
Errors in arithmetic, 142 ff.
Errors in spelling, 194 ff.

Fordyce, Charles, 182.
Freeman, F. N., 198, 199, 212, 214, 215, 233.
Freeman's Handwriting Scale, 212.
Frostic, F. W., 244.

Geography: Courtis's Standard Tests in Geography, 256 ff.; Hahn-Lackey Geography Scale, 258 ff.; Boston Geography Tests, 264; Buckingham's Geography Test, 265; Starch's Geography Tests, Series A, 265; Witham's Standard Geography Tests, 265; Branom and Reavis Completion Test for Geography, 265.
Gist, Arthur S., 142.
Gray, C. T., 216.
Gray, W. S., 67.
Gray's Oral Reading Test, 39 ff.
Gray's Score Card for Handwriting, 216 ff.
Gray's Silent Reading Tests, 67.

Haggerty, M. E., 49.
Hahn, H. H., 258.
Hahn-Lackey Geography Scale, 258 ff.
Handwriting: Ayres's Scale, "Gettysburg Edition," 208 ff.; Freeman's Scale, 212; Gray's Score Card, 216 ff.; measurement of rate, 203 ff., measurement of quality, 207; standards, 219 ff.
Harlan's Test of Information in American History, 262 ff.
Harvard-Newton Composition Scale, 244.
Hillegas Composition Scale, 244.
History: Harlan's Test of Information in American History, 262 ff.; Buckingham's History Test, 265; Bell and McCollum History Test, 266; Starch's American History Tests, Series A, 266.
Hollingworth, Leta S., 196, 197.

Johnson, F. W., 4.
Jones, N. F., 176, 192.
Judd, C. H., 67, 73, 76, 84, 225.

Kallom, Arthur W., 109, 144, 197.
Kelly, F. J., 3, 9, 278.
King, W. I., 105.
Koos, L. V., 222.

Lackey, E. H., 258.
Language and Grammar: Charters's Diagnostic Test in Language and Grammar, 245; Starch's Punctuation Scale, 248–49; Boston Copying Test, 249.
Lewis, E. E., 223.
Lull, H. G., 201.

Median, calculation of, 29, 35, 102 ff.
Monroe, Walter S., 109.
Monroe's Diagnostic Tests, 109.
Monroe's Standardized Reasoning Tests, 154 ff.
Monroe's Standardized Silent Reading Tests, 22 ff., 43 ff., 51 ff.
Monroe's Timed-Sentence Spelling Test, 185, 189.

Nassau County Supplement, 243.
Nutt, H. W., 225.

Otis, A. S., 180, 181.

Race, Henrietta V., 49.

Reading: Monroe's Standardized Silent Reading Tests, 22 ff.; Courtis's Silent Reading Test No. 2, 29 ff.; Thorndike's Visual Vocabulary Scale, 36 ff., 49; Gray's Oral Reading Test, 39 ff., 49–51; Thorndike's Scale, Alpha 2, for the understanding of sentences, 54.
Reavis, C. W., 233.
Rugg, H. O., 262.
Rural Schools, use of test in, 25, 32, 88 ff., 91, 99–100, 137.

Scores, good arrangement of, 26.
Sears, J. B., 194.
Smith, James H., 151.
Spelling: Ayres's Measuring Scale for Ability in Spelling, 176 ff.; Monroe's Timed-Sentence Spelling Test, 185 ff.
Spelling demons, 192.
Spelling games, 193.
Standards: Monroe's Standardized Silent Reading Tests, 44; Courtis's Silent Reading Test No. 2, 46; Thorndike's Visual Vocabulary Scale, 49; Gray's Oral Reading Test, 51; Courtis's Standard Research Tests, Series B, 108; Monroe's Diagnostic Tests, 116; Monroe's Standardized Reasoning Tests, 157; Ayres's Measuring Scale for Ability in Spelling, 175 ff., 187–89; Monroe's Timed-Sentence Spelling Test, 185, 189; Handwriting, 219 ff., 231; Willing's Scale for Measuring Written Composition, 243; Starch's Punctuation Scale, 249; Boston Copying Test, 252; Charters's Diagnostic Test in Language and Grammar, 248; Hahn-Lackey Geography Scale, 261; Harlan's Test of Information in American History, 263.
Standards, necessity of, 18.
Starch, Daniel, 8, 9, 174, 181, 248.
Starch's American History Tests, Series A, 266.
Starch's Geography Tests, Series A, 265.
Starch's Punctuation Scale, 248–49.
Stone, C. R., 230.
Stone, C. W., 109, 159, 173.
Studebaker's Economy Practice Exercises, 136.

Terman, L. M., 71.
Thorndike, E. L., 54, 56, 57, 105, 180, 181.
Thorndike's Extension of the Hillegas Composition Scale, 244.
Thorndike's Handwriting Scale, 208.
Thorndike's Scale, Alpha 2, for the Understanding of Sentences, 54.
Thorndike's Visual Vocabulary Scale, 36 ff.

Uhl, W. L., 92, 149.

Vocabulary: Thorndike's Visual Vocabulary Scale, 36 ff.
Vocabulary in arithmetic, 163, 165.

Wasson, Alfred W., 94.
Willing, M. H., 235.
Willing's Scale for Measuring Written Composition, 235 ff.
Wilson, H. B., 229.
Wilson, G. M., 229.
Witham's Standard Geography Tests, 265.

Yawberg, A. G., 269.

Zirbes, Laura, 79, 86 ff., 91.

RIVERSIDE TEXTBOOKS IN EDUCATION

HEALTHFUL SCHOOLS: HOW TO BUILD, EQUIP, AND MAINTAIN THEM.

By May Ayres, Jesse F. Williams, M. D., Professor of Hygiene and Physical Education, University of Cincinnati, and Thomas D. Wood, A. M., M. D., Professor of Physical Education, Teachers College, Columbia University. $1.50 *net.* Postpaid.

TEACHING LITERATURE IN THE GRAMMAR GRADES AND HIGH SCHOOL.

By Emma M. Bolenius, formerly Instructor in English, Central Commercial and Manual Training High School, Newark, N. J. $1.50 *net.* Postpaid.

PUBLIC EDUCATION IN THE UNITED STATES.

By Ellwood P. Cubberley. *In preparation.*

PUBLIC SCHOOL ADMINISTRATION.

By E. P. Cubberley. $2.00 *net.* Postpaid.

RURAL LIFE AND EDUCATION.

By Ellwood P. Cubberley, Dean of the School of Education, Leland Stanford Junior University. $1.60 *net.* Postpaid.

EVOLUTION OF THE EDUCATIONAL IDEAL.

By Mabel I. Emerson, First Assistant in Charge of the George Bancroft School, Boston. $1.20 *net.* Postpaid.

EXPERIMENTAL EDUCATION.

By F. N. Freeman. $1.50 *net.* Postpaid.

HOW CHILDREN LEARN.

By F. N. Freeman. $1.60 *net.* Postpaid.

THE PSYCHOLOGY OF THE COMMON BRANCHES.

By F. N. Freeman, Assistant Professor of Educational Psychology, University of Chicago. $1.50 *net.* Postpaid.

HEALTH WORK IN THE SCHOOLS.

By E. B. Hoag, M. D., Medical Director, Long Beach City Schools, California, and L. M. Terman. $1.75 *net.* Postpaid.

HOW TO TEACH THE FUNDAMENTAL SUBJECTS.

By C. N. Kendall, Commissioner of Education for New Jersey, and G. A. Mirick, formerly Deputy Commissioner of Education for New Jersey. $1.50 *net.* Postpaid.

HOW TO TEACH THE SPECIAL SUBJECTS.

By C. N. Kendall and G. A. Mirick. $1.60 *net.* Postpaid.

EDUCATIONAL TESTS AND MEASUREMENTS.

By W. S. Monroe, Director of the Bureau of Coöperative Research, University of Indiana; J. C. DeVoss, Associate Professor of Psychology and Philosophy, Kansas State Normal School; and F. J. Kelly, Dean of the School of Education, University of Kansas. $1.60 *net.* Postpaid.

MEASURING THE RESULTS OF TEACHING.

By W. S. Monroe. $1.60 *net.* Postpaid.

DISCIPLINE AS A SCHOOL PROBLEM.

By A. C. Perry, Jr., District Superintendent of Schools, New York City. $1.50 *net*. Postpaid.

STATISTICAL METHODS APPLIED TO EDUCATION.

By H. O. Rugg, Assistant Professor of Education, University of Chicago. $2.00 *net*. Postpaid.

CLASSROOM ORGANIZATION AND CONTROL.

By J. B. Sears, Associate Professor of Education, Leland Stanford Junior University. $1.75 *net*. Postpaid.

AN INTRODUCTION TO EDUCATIONAL SOCIOLOGY.

By Walter R. Smith, Professor of Sociology and Economics, Kansas State Normal School, Emporia, Kansas. $1.75 *net*. Postpaid.

THE HYGIENE OF THE SCHOOL CHILD.

By L. M. Terman, Professor of Education, Leland Stanford Junior University. $1.75 *net*. Postpaid.

THE INTELLIGENCE OF SCHOOL CHILDREN.

By L. M. Terman. *In preparation.*

THE MEASUREMENT OF INTELLIGENCE.

By L. M. Terman. $1.75 *net*. Postpaid.

Test Material for the Measurement of Intelligence. 60 cents *net*. Postpaid. Record Booklets (in packages of twenty-five). $2.00 *net*, a package. Postpaid.

THE TEACHING OF SCIENCE IN THE ELEMENTARY SCHOOL.

By Gilbert H. Trafton, Instructor in Science, State Normal School, Mankato, Minnesota. $1.30 *net*. Postpaid.

AN INTRODUCTION TO CHILD PSYCHOLOGY.

By Charles W. Waddle, Ph. D., Supervisor of Practice Teaching, Los Angeles State Normal School. $1.50 *net*. Postpaid.

TEACHING IN RURAL SCHOOLS.

By T. J. Woofter, Dean of the School of Education, University of Georgia. $1.40 *net*. Postpaid.

Secondary Education Division

PRINCIPLES OF SECONDARY EDUCATION.

By Alexander Inglis, Assistant Professor of Education, Harvard University. $2.75.

PROBLEMS OF SECONDARY EDUCATION.

By David Snedden, Professor of Education, Teachers College, Columbia University. $1.50 *net*. Postpaid.

THE TEACHING OF ENGLISH IN THE SECONDARY SCHOOLS.

By Charles Swain Thomas, Head of the English Department, Newton High School, Newton, Mass. $1.60 *net*. Postpaid.

HOUGHTON MIFFLIN COMPANY
BOSTON NEW YORK CHICAGO

1725 b

ON VOCATIONAL PREPARATION

SCHOOL CREDIT FOR HOME WORK By L. R. Alderman

A successful plan for the coöperation of school and home. $1.25. Postpaid.

THE VOCATIONAL GUIDANCE OF YOUTH
By Meyer Bloomfield

A monograph by the former Director of the Vocation Bureau of Boston. 75 cents. Postpaid.

YOUTH, SCHOOL, AND VOCATION By Meyer Bloomfield

A first-hand presentation of the meaning and work of the vocational guidance movement. $1.35. Postpaid.

CHOOSING A VOCATION By Frank Parsons

This book is an indispensable manual for every vocational counselor. $1.10. Postpaid

THE PROBLEM OF VOCATIONAL EDUCATION
By David Snedden

The author is the Professor of Education, Teachers College, and one of the leaders in the movement for the closer adaptation of public schools to the actual needs of youth. 40 cents. Postpaid.

PREVOCATIONAL EDUCATION
By Frank M. Leavitt and Edith Brown

The first authoritative book to tell how the public schools may prepare pupils to select wisely the work to which they are best adapted. $1.30. Postpaid.

THE PEOPLE'S SCHOOL By Ruth Mary Weeks

A statement regarding the vocational training movement in this country and abroad. 75 cents. Postpaid.

VOCATIONS FOR GIRLS By Mary A. Laselle and Katherine Wiley

Information as to conditions of work and the opportunities in the more common vocations open to girls with only a high-school education. $1.00. Postpaid.

THE HOME SCHOOL By Ada Wilson Trowbridge

An account of a unique and successful experiment in teaching practical household economics. 75 cents. Postpaid.

VOCATIONAL EDUCATION
By David Snedden, Ruth Mary Weeks, and Ellwood P. Cubberley

A combination of three volumes from the *Riverside Educational Monographs* treating different phases of vocational education, — theory, administration, and practice. $1.35. Postpaid.

INDUSTRIAL EDUCATION: Its Problems, Methods, and Dangers By Albert H. Leake

A study and criticism of the opportunities provided for the education of the industrial worker. $1.40. Postpaid.

ESTABLISHING INDUSTRIAL SCHOOLS
By Harry Bradley Smith

A practical discussion of the steps to be taken in establishing industrial schools. 75 cents. Postpaid.

HOUGHTON MIFFLIN COMPANY
BOSTON NEW YORK CHICAGO

1718

THE HOUGHTON MIFFLIN PROFESSIONAL LIBRARY

For Teachers and Students of Education

PRACTICAL ASPECTS OF EDUCATION

	Postpaid
Charters's Teaching the Common Branches	$1.50
Nolan's The Teaching of Agriculture	1.30
Earhart's Types of Teaching	1.40
Wilson's The Motivation of School Work	1.40
Leavitt and Brown's Prevocational Education in the Public Schools	1.30
Hall's The Question as a Factor in Teaching	1.30
Kready's A Study of Fairy Tales	1.40
Bryant's How to Tell Stories to Children	1.10
Cabot's Ethics for Children	1.50
Brownlee's Character Building in School	1.25
A Course in Citizenship and Patriotism	1.50
Bloomfield's Youth, School, and Vocation	1.35
Colby's Literature and Life in School	1.35
The Kindergarten	1.50
Fulmer's The Use of the Kindergarten Gifts	1.30
Bates's Talks on Teaching Literature	1.50

THEORY AND PRINCIPLES OF EDUCATION

McMurry's (F. M.) How to Study and Teaching How to Study	1.40
McMurry's (C. A.) Conflicting Principles in Teaching	1.40
Woodley's The Profession of Teaching	1.40
Bobbitt's The Curriculum	1.50
Kirkpatrick's Fundamentals of Sociology	1.40
Kirkpatrick's The Individual in the Making	1.40
Ruediger's The Principles of Education	1.50
Hanus's Beginnings in Industrial Education	1.25
O'Shea's Social Development and Education	2.00
Tyler's Growth and Education	1.50
Henderson's Education and the Larger Life	1.50
Chancellor's A Theory of Motives, Ideals and Values in Education	2.00

HOUGHTON MIFFLIN COMPANY
BOSTON NEW YORK CHICAGO

1823

150977
Educat.
Teach.
M.

Author Monroe, Walter Scott
Title Measuring the results of teaching.

University of Toronto
Library

DO NOT
REMOVE
THE
CARD
FROM
THIS
POCKET

Acme Library Card Pocket
Under Pat. "Ref. Index File"
Made by LIBRARY BUREAU